

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/001543

International filing date: 16 February 2005 (16.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE

Number: 10 2004 013 988.1

Filing date: 19 March 2004 (19.03.2004)

Date of receipt at the International Bureau: 29 March 2005 (29.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

BUNDESREPUBLIK DEUTSCHLAND

16. 02. 2005



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 10 2004 013 988.1

Anmeldetag: 19. März 2004

Anmelder/Inhaber: Henkel Kommanditgesellschaft auf Aktien,
40589 Düsseldorf/DE

Bezeichnung: Der Faktor RecA aus *Bacillus licheniformis* und
recA-inaktivierte Sicherheitsstämme für die bio-
technologische Produktion

IPC: C 07 K, C 12 N

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 25. November 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

A handwritten signature in black ink, appearing to read "Brosig", is placed over the typed name "Der Präsident". The signature is written in a cursive, flowing style.

Brosig

Henkel KGaA
Dr. Stock/KF
19.03.2004

AZ 10 2004 013 988,1
AT 19.03.04


Patentanmeldung

H 06291

Der Faktor RecA aus *Bacillus licheniformis* und recA-inaktivierte Sicherheitsstämme für die biotechnologische Produktion

Die vorliegende Erfindung betrifft den Faktor RecA aus *Bacillus licheniformis* DSM 13 sowie Mikroorganismen als Sicherheitsstämme für die biotechnologische Produktion, die dadurch gekennzeichnet sind, daß sie funktionelle Deletionen in dem zugehörigen Gen recA aufweisen. Ferner steht RecA damit für weitere molekularbiologische Ansätze zur Verfügung.

Die vorliegende Erfindung liegt auf dem Gebiet der Biotechnologie, insbesondere der Herstellung von Wertstoffen durch Fermentation von Mikroorganismen, die zur Bildung der interessierenden Wertstoffe in der Lage sind. Hierzu zählen beispielsweise die Herstellung niedermolekularer Verbindungen, etwa von Nahrungsmittelergänzungsstoffen oder pharmazeutisch relevanten Verbindungen, oder von Proteinen, für welche aufgrund ihrer Diversität wiederum ein großes technisches Einsatzgebiet besteht. Im ersten Fall werden die Stoffwechseleigenschaften der betreffenden Mikroorganismen zur Herstellung der Wertstoffe ausgenutzt und/oder verändert; im zweiten Fall werden Zellen eingesetzt, die die Gene der interessierenden Proteine exprimieren. In beiden Fällen handelt es sich zumeist also um gentechnisch veränderte Organismen (GVO).

Zur Fermentation von Mikroorganismen besteht ein reichhaltiger Stand der Technik, insbesondere auch im großtechnischen Maßstab; er reicht von der Optimierung der betreffenden Stämme hinsichtlich der Bildungsrate und der Nährstoffausnutzung über die technische Gestaltung der Fermenter bis hin zur Gewinnung der Wertstoffe aus den betreffenden Zellen selbst und/oder dem Fermentationsmedium. Hierfür kommen sowohl genetische und mikrobiologische als auch verfahrenstechnische und biochemische Ansätze zu tragen. Ziel der vorliegenden Erfindung ist es, diesen Prozeß hinsichtlich der sicherheitsrelevanten Aspekte der eingesetzten Mikroorganismen zu verbessern, und zwar auf der Ebene der genetischen Eigenschaften der betrachteten Stämme.

Dies steht vor dem Hintergrund, daß die Verwendung gentechnisch veränderter Organismen im allgemeinen strengen gesetzlichen Richtlinien bezüglich der biologischen Sicherheit unterliegt. In den meisten Ländern sind die Betreiber von Anlagen mit GVO gehalten, dafür zu sorgen, daß nach Möglichkeit keine GVO in die Umgebung gelangen. Zusätzlich sollen für die Produktion verwendete GVO Eigenschaften aufweisen, die es ihnen – falls sie doch in die Umgebung gelangen sollten – erschweren oder je nach Gefahrenpotential sogar unmöglich machen sollen, sich dort zu vermehren („Containment“-Konzept).

Dem Übersichtsartikel „Suicidal genetic elements and their use in biological containment of bacteria“ von S.Molin et al. (Annu. Rev. Microbiol., 1993, Band 47, Seiten 139 bis 166) zufolge wird dabei zwischen den beiden grundsätzlichen Strategien differenziert, als „aktive“ Komponenten, kontrollierte Suizidsysteme in die Zellen einzuführen oder über „passive“ Systeme die Zelleigenschaften so zu verändern, daß ihre Überlebenschancen unter Stressbedingungen sinken. Die zweite, für die vorliegende Anmeldung relevante Strategie wird darin auch als „Disablement approach“ bezeichnet.

Stämme von GVO mit einem verminderten Risiko für Mensch und Umwelt im Falle einer unbeabsichtigten Freisetzung werden als Sicherheitsstämme bezeichnet. Je nach den grundsätzlichen Eigenschaften der Mikroorganismen werden zunehmend mehrere Eigenschaften gefordert, von denen jede für sich einen Sicherheitsaspekt darstellt. Somit ist es vorteilhaft, verschiedene Instrumente zur Erzeugung von Sicherheitsstämmen zur Verfügung zu haben. Hiervon sind einige „passive“ Systeme bereits im Stand der Technik beschrieben.

So betrifft die Anmeldung EP 369817 A1 Bacillus-Stämme, insbesondere *B. subtilis*, zur Herstellung und Sekretion von Proteinen, bei denen die Gene für extra- und intrazelluläre Proteasen, nämlich *epr*, *rp-I*, *rp-II*, *isp-1*, *apr* und/oder *npr* durch Punktmutationen oder Insertionen inaktiver Genkopien funktionell inaktiviert worden sind. Der Sinn dieser gentechnischen Veränderungen besteht darin, die für die mit diesen Stämmen hergestellten, interessierenden Proteine schädlichen Protease-Aktivitäten zu minimieren. Die betreffenden Stämme können zusätzlich über Mutationen verfügen, die die Sporulation und damit eine Bildung ebenso schädlicher Sporulationsproteasen verhindern. Hierunter wird das in Phase Null der Sporulation (siehe unten) von *B. subtilis*

aktive Gen spoOA genannt, um durch dessen Inaktivierung die mit der Sporulation verbundene Bildung intrazellulärer Proteasen zu verhindern.

Die Anmeldung WO 92/16642 A1 verfolgt den gleichen Lösungsansatz: sie offenbart, daß durch Inaktivierung der Proteasegene apr, npr, isp-1, epr, bpr, rsp und mpr von *Bacillus* ein Großteil der extrazellulären Protease-Aktivität ausgeschaltet wird, und lehrt, daß dies durch die Inaktivierung des neu beschriebenen Gens vpr für die Rest-Protease III noch verbessert werden kann. Auch hier wird auf die Möglichkeit der Inaktivierung von spoOA hingewiesen, um die Bildung intrazellulärer Proteasen zu verhindern.

Bei der Sporulation von grampositiven Bakterien handelt es sich um einen Entwicklungsprozeß zur Bildung von Dauerformen, den sogenannten Sporen, zum Überdauern widriger Umwelteinflüsse. Sie wird über eine komplexe Regulationskaskade mit vermutlich mehr als 100 Genen und unter Beteiligung von speziellen Sigma-Faktoren gesteuert. Den Zusammenhang dieses Prozesses mit dem Zellzyklus von *B. subtilis* beschreibt beispielsweise die Publikation „Cell cycle and sporulation in *Bacillus subtilis*“ (1998) von P.A. Levin und A.D. Grossmann in *Curr. Opin. Microbiol.*, Band 1, Seiten 630 bis 635. Hier wird vor allem der Transkriptionsfaktor Spo0A als Kontrollelement zum Einleiten der Sporulation dargestellt. Die sequentielle Aktivierung der phasenspezifischen Gene durch verschiedene Sigmafaktoren faßt beispielsweise der Übersichtsartikel „Control of sigma factor activity during *Bacillus subtilis* sporulation“ (1999) von L. Kroos et al. in *Mol. Microbiol.*, Band 31, Seiten 1285 bis 1294 zusammen. Bei diesem Vorgang werden ihrer Reihenfolge nach die aufeinanderfolgenden Stadien Null und dann I bis VII beobachtet. Diese Numerierung findet sich auch in den Bezeichnungen der beteiligten Gene und Faktoren wieder.

Die Anmeldung EP 492274 A2 offenbart, daß im Stand der Technik bereits über unspezifische Mutagenese die Inaktivierung von Sporulationsgenen gelungen sei, wodurch asporogene Mutanten (spo-Minus-Phänotyp) erhalten worden seien.

EP 492274 A2 selbst beschreibt einen durch gezielte Mutagenese in dem frühen Sporulationsgen spoII:D behandelten *B. subtilis*-Stamm, der mit einer Reversionsrate von weniger als 10^{-8} praktisch nicht mehr in der Lage ist, Sporen zu bilden. Diese Anmeldung lehrt, diesen Stamm erst nach Inaktivierung der weiteren Gene leu (für die Leucin-Synthese), pyrD1 (für die Uracil-Synthese), apr und npr für die Herstellung von

Wertstoffen für die biotechnologische Produktion zu verwenden, weil hiermit Vorteile in der Produktion sowie Sicherheitsaspekte verbunden seien.

Die Anmeldung WO 97/03185 A1 befaßt sich ebenfalls mit der Inaktivierung der Sporulationsfähigkeit von *Bacillus*-Spezies mit Ausnahme von *B. subtilis* und Verwendung dieser Stämme zur biotechnologischen Herstellung von Wertstoffen. Dieser Anmeldung zufolge soll das frühe, für den Sigmafaktor F codierende Gen *spolIAC* funktionell inaktiviert werden, vorteilhafterweise in Kombination mit Deletionen in Genen der ebenfalls früh aktivierten Sporulationsgruppen *spo2*, *spo3*. Hierfür wird eine irreversible Inaktivierung der betreffenden chromosomalen Abschnitte für *spolIAC* beschrieben.

Die Anmeldung WO 02/097064 A1 (EP 1391502 A1) betrifft Mikroorganismen, bei denen Gene aus den Stadien II, III, IV oder V der Sporulation delétiert oder inaktiviert worden sind. Hierbei handelt es sich um die Gene *sigE*, *sigF*, *spolIE*, *spolISB* und *sigG* von *B. subtilis*, die innerhalb des Genorts von *spolVCB* bis *spolIIC* von *B. subtilis* liegen. Dieser kann anhand der Datenbank SubtiList (zugänglich über <http://genolist.pasteur.fr/SubtiList/genome.cgi>) auf den Bereich der Positionen von ca. 2.642.000 kb bis ca. 2.700.000 kb des inzwischen bekannten Gesamtgenoms von *B. subtilis* eingegrenzt werden. Dieser Anmeldung hatte die Aufgabe zugrunde gelegen, überflüssige oder schädliche Aktivitäten von *Bacillus*-Stämmen auszuschalten, um die biotechnologische Produktion zu verbessern. Durch die genannten Modifikationen der mittleren bis späten Sporulationsgene würde bei Einsatz der betreffenden Stämme für die biotechnologische Produktion die Sporenbildung unterdrückt; dies wirke sich vorteilhaft auf die Nährstoff- und Energieverwertung aus; gleichzeitig könne die Dauer der Fermentation erhöht werden, um darüber die Gesamtausbeute an interessierendem Wertstoff zu erhöhen.

Der Übergang grampositiver Bakterien in die Dauerform der Spore wird also durch ungünstige Umweltbedingungen ausgelöst. Genau dies dürfte auch dann geschehen, wenn Bakterien aus den optimalen Wachstumsbedingungen der Fermentation unbeabsichtigt aus der Anlage austreten und in die Umgebung gelangen sollten. Demgegenüber ist, wie soeben dargelegt, über die Unterbindung der Fähigkeit zur Sporulation zur Erzeugung sicherer GVO bislang erst wenig nachgedacht worden. Der einschlägige Stand der Technik scheint lediglich nahezulegen, daß die Sporulation grampositiver Bakterien (a) wegen der damit verbundenen Proteaseaktivitäten und/oder

(b) zur Verlängerung der Fermentationsdauer frühzeitig und vollständig unterbunden werden sollte, um letztlich die Fermentationsausbeute der so erhaltenen asporogenen Stämme zu steigern. Zur Verfolgung von Sicherheitsaspekten werden dagegen nur mehrere, zusätzlich einzuführende Mutationen offenbart.

Das für den in Prokaryonten beschriebenen Faktor RecA codierende Gen *recA* ist in der Molekularbiologie sehr bekannt, bislang insbesondere jedoch aus einem anderen Zusammenhang als dem der Herstellung von Sicherheitsstämmen. Dieser Faktor bindet spezifisch und kooperativ an einzelsträngige DNA und sorgt unter ATP-Hydrolyse für eine teilweise Entwindung doppelsträngiger DNA. Dieser Vorgang ermöglicht den genetischen Vorgang der Rekombination, das heißt den Strangtausch zwischen ähnlichen DNA-Molekülen. So ist es in der Molekularbiologie ein übliches Vorgehen, das *recA*-Gen, dadurch zu verwenden, daß es über ein entsprechendes genetisches Konstrukt mit einer defekten *recA*-Kopie inaktiviert und somit ein *recA*-Minus-Phänotyp erzeugt wird, welcher nicht mehr zur Rekombination in der Lage ist. Beispielsweise nach dem Patent US 4713337 werden durch Crossing-over erzeugte Deletionsmutanten durch anschließende Inaktivierung von *recA* genetisch stabilisiert.

Eine biochemische Beschreibung des RecA aus *Escherichia coli* liefert beispielsweise die Publikation „C-terminal deletions of the *Escherichia coli* RecA“ von S.L.Lusetti et al. (2003; J. Biol. Chem., Band 278, Heft 18, Seiten 16372-16380). Daraus geht hervor, daß insbesondere der C-Terminus dieses Moleküls die Einzelstrangbindung vermittelt und entsprechende Deletionsmutanten in diesem Bereich neben anderen biochemischen Eigenschaften eine erhöhte Mitomycin-Sensitivität aufweisen. Mitomycin ist dafür bekannt, daß es die DNA-Synthese beeinträchtigt und dadurch bakterizid wirkt. Der N-terminale Bereich ist dagegen stärker an der Bindung von DNA-Doppelsträngen beteiligt.

Der bereits zitierte Übersichtsartikel von S.Molin et al. verweist ebenfalls auf Arbeiten, bei denen eine *recA*-Minus-Mutation als Markgen für den gramnegativen *Escherichia coli* verwendet wird. Es wird die Vermutung geäußert, diese Mutation allein könne schon ausreichen, um jegliches Umweltrisiko durch diesen Stamm auszuschließen. Andererseits werden zwei Nachteile dieses Ansatzes diskutiert, nämlich zum einen sei es technisch schwierig, diese Mutanten herzustellen und zum anderen würden die betreffenden Stämme auch in ihrem erwünschten Einsatz („short-term competitive properties“) derart behindert sein, daß man andere die Lebensfähigkeit beschränkende und im betreffenden

Artikel erwähnte Mutationen vorziehen würde. Auch in Kombination mit dem grundsätzlich anderen Ansatz zur Erzeugung von Sicherheitsstämmen, nämlich dem Einführen von Suizidsystemen, habe sich die Inaktivierung von RecA als nachteilhaft herausgestellt.

Die Publikation „Freisetzung gentechnisch veränderter Bakterien“ von Selbitschka et al. (2003; Biologie in unserer Zeit, Band 33, Heft 3, Seiten 162-175) beschreibt die Freisetzung von gramnegativen, mit einem Luciferase-Gen modifizierten Bakterien der Spezies *Sinorhizobium meliloti* in einem mehrjährigen Freilandversuch. Sie trugen zusätzlich eine Inaktivierung des *recA*-Gens, welche dazu geführt hat, daß Zellen dieser Klone unter natürlichen Bedingungen letztlich nicht überleben konnten.

K.-D. Wittchen hat im Zuge seiner bei der Westfälischen Wilhelms-Universität Münster eingereichten Dissertation (1995) mit dem Titel „Entwicklung eines Sicherheitsstammes von *Bacillus megaterium* DSM 319 und molekulargenetische Charakterisierung des Gens für die extrazelluläre neutrale Metalloprotease (*nprM*)“ einen Stamm des grampositiven *B. megaterium* erzeugt, der nach gezielter Gendisruption Deletionen in der im Titel genannten neutralen Metalloprotease, der der Isopropylmalat-Dehydrogenase, und eines nicht näher bezeichneten SpolV-Proteins enthält. Anschließend wurde an diesem Stamm eine *recA*-Mutation vorgenommen, welche in bezug auf die UV-Sensitivität des Stammes im Vergleich zum Wildtyp allerdings keine wesentlichen Unterschiede verursachte, wohl aber bei Wachstum auf Mitomycin C-haltigen Agarplatten. Diese Vierfachmutante wurde als Sicherheitsstamm vorgeschlagen, ohne jedoch dessen Realisierbarkeit oder sogar die konkreten Auswirkungen dieser Modifikationen auf einen Produktionsprozeß mit entsprechend modifizierten Bakterienstämmen zu untersuchen.

Die in derselben Arbeitsgruppe angefertigte Diplomarbeit von H. Nahrstedt (2000) mit dem Titel „Molekulargenetische Charakterisierung des *recA*-Gens von *Bacillus megaterium* DSM 319 und Konstruktion einer Deletionsmutante“ schlägt folgende, nebeneinander vorliegende vier Mutationen in zum Teil Gruppen von Genen vor: *recA*-Minus, Protease-Minus, Leucin-Auxotrophie und Sporulations-Defizienz. Es wird diskutiert, eine *recA*-Defizienz als eine Sicherheitsmarke neben anderen in Produktionsstämme einzuführen, weil sich dadurch zum einen unerwünschte Rekombinationsprozesse unterdrücken ließen und zum anderen die betreffenden Mutanten eine erhöhte Sensitivität gegenüber DNA-schädigenden Agenzien aufweisen,

das heißt in der Umwelt eine geringere Überlebenschance besitzen sollten. Auch dieser Vorschlag wurde nicht weiterverfolgt.

Den Stand der Technik zu RecA kann man dahingehend zusammenfassen, daß dieses Protein bisher überwiegend aus genetischen Zusammenhängen bekannt ist. Der Einsatz dieses Faktors bzw. die Inaktivierung des betreffenden Gens zur Erzeugung von Sicherheitsstämmen von GVO wird aufgrund seiner physiologischen Bedeutung bislang eher abgelehnt. Erfolgreiche beschreibbare Beispiele hierfür sind lediglich recA-Minus-Mutanten der gramnegativen Spezies *Sinorhizobium meliloti* und des grampositiven *Bacillus megaterium*, letztere jeweils nur in Kombination mit drei weiteren sicherheitsrelevanten Mutationen.

Insgesamt bleibt festzuhalten, daß zur Herstellung von Sicherheitsstämmen für die biotechnologische Produktion verschiedene alternative genetische Systeme nach einem „passiven“ Wirkmechanismus etabliert sind, wie die Inaktivierung von Proteasegenen, die Ausschaltung von verschiedenen Stoffwechselgenen zur Erzeugung von Aminosäure- oder Nukleobasenauxotrophien. Bei sporenbildenden grampositiven Bakterien ist die Unterbindung der Sporulation beschrieben, insbesondere zu einem frühen Zeitpunkt, vor allem aber um zusätzliche Vorteile in der Fermentation zu erzielen.

Es stellte sich somit die Aufgabe, ein weiteres geeignetes Sicherheitssystem für gentechnisch veränderte grampositive Bakterien zu entwickeln, wofür als Grundlage zunächst ein geeigneter Faktor und/oder ein geeignetes Gen zu identifizieren waren.

Einen Teilespekt dieser Aufgabe stellte nach Feststellung der grundsätzlichen Eignung eines solchen Systems die Isolierung eines hierfür verwendbaren genetischen Elements, eventuell eines Gens, und der Aminosäuresequenz eines hiervon gegebenenfalls codierten Faktors dar, um dieses System entsprechenden molekularbiologischen Konstruktionen für den Einsatz in Produktionsstämmen zugänglich zu machen, insbesondere in Kombination mit einem oder mehreren weiteren der Sicherheit dienenden Regulationsmechanismen.

Ein weiterer Teilespekt der Aufgabe bestand darin, daß dieses System mit anderen Sicherheitssystemen kombinierbar sein sollte.

Somit lag eine Teilaufgabe darin, ein weiteres derartiges, hiermit zu kombinierendes Sicherheitssystem zu definieren, vorzugsweise eines, das neben diesen beiden Systemen keine weiteren Mutationen erforderlich machen würde. Mit anderen Worten: maximal diese zwei Mutationen sollten ausreichen, um einen grampositiven Sicherheitsstamm zu erzeugen, der weitgehende Anforderungen an die Verminderung der Lebensfähigkeit in der Umwelt erfüllen, das heißt zu einer minimalen Reversionsrate führen sollte. Denn eine niedrigere Zahl als vier nebeneinander wirksame Systeme bedeutet einen zunehmend geringeren Arbeitsaufwand zur Herstellung dieser Stämme.

Ein Nebenaspekt dieser Aufgabe bestand darin, ein derartiges Sicherheitssystem zu finden, das nicht so speziell ist, daß es nicht auch in anderen molekularbiologischen Ansätzen Verwendung finden könnte.

Diese Aufgabe wird durch den Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz mindestens zu 96% identisch ist, beziehungsweise durch die für einen Faktor RecA codierende Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist, gelöst.

Die in SEQ ID NO. 2 und 1 angegebenen Aminosäure- und Nukleotidsequenzen sind die für RecA. Dabei codieren alle Positionen von 1 bis 1047 für das Protein; die letzten drei stellen dabei das Stop-Codon dar. Sie werden als Gen und Protein recA beziehungsweise RecA bezeichnet. Sie stammen aus dem bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig (<http://www.dsmz.de>) unter der Nummer DSM 13 hinterlegten Stamm *Bacillus licheniformis*. Erfnungsgemäße Lösungen der Aufgabe stellen alle Faktoren beziehungsweise Nukleinsäuren dar, die hierzu eine hinreichend Homologie aufweisen, wie sie mit den jeweiligen Prozentangaben definiert ist.

Als nächster Stand der Technik kann der entsprechende Faktor aus *B. amyloliquefaciens* angesehen werden. Die zugehörigen DNA- und Aminosäuresequenzen sind in der Datenbank NCBI der National Institutes of Health der USA (<http://www.ncbi.nlm.nih.gov>) unter der Zugangsnummer AJ515542 veröffentlicht. RecA aus *B. licheniformis* DSM 13 weist hierzu auf Aminosäureebene eine Homologie von 94,0% Identität und auf Nukleinsäureebene eine Übereinstimmung von 81,2% Identität auf. Beide Vergleiche

gehen aus den Alignments der Figuren 1 und 2 hervor, wo die Sequenzen von *B. amyloliquefaciens* jeweils in der zweiten Zeile dargestellt sind.

Als nächstähnliche Enzyme wurden RecA aus *B. subtilis* und RecE aus *B. subtilis* mit jeweils 93,4% Identität ermittelt. Sie weisen auf DNA-Ebene Homologiewerte von 81,0% beziehungsweise 81,2% Identität auf. Sie sind ebenfalls in der NCBI-Datenbank, und zwar unter den Eintragungsnummern Z99112 (Region 161035 bis 162078) beziehungsweise X52132 veröffentlicht. Die Aminosäure- und DNA-Vergleiche mit diesen Faktoren sind ebenfalls in den Figuren 1 und 2 (jeweils Zeilen 3 beziehungsweise 4) dargestellt.

Diese hohen, beanspruchten Homologiewerte um den konkret hier beschriebenen Faktor lassen erwarten, daß derselbe Faktor RecA besonders in verwandten Stämmen oder Spezies, wahrscheinlich aber auch in weniger verwandten Spezies, wahrscheinlich sogar gramnegativen Organismen die zugehörige Funktion übernimmt. Diese liegt erfindungsgemäß in der eingangs erläuterten DNA-Einzelstrangbindung und der damit verbundenen Rolle für Rekombinationsvorgänge von Nukleinsäuren. Vergleichbare Wirkungen sollten auch mit Deletionen des recA-Gens verbunden sein, nämlich die Unterbindung von DNA-Rekombinationen und eine dadurch verminderte Lebensfähigkeit. Gleichzeitig sollte ein recA-Gen aus dem einen Stamm geeignet sein, diese Funktion in einem anderen zu übernehmen; dies wird mit zunehmender Ähnlichkeit zunehmend besser gelingen. Hierdurch wird der Einsatz des betreffenden Gens für die Herstellung von Deletionsmutanten der verschiedensten grampositiven Mikroorganismen möglich.

Dadurch wird für RecA und vor allem für das zugehörige Gen recA ein weites technologisches und kommerziell relevantes Gebiet eröffnet, nämlich die Herstellung von Wertstoffen durch Fermentation von genetisch veränderten grampositiven Bakterien. Sie können über Mutationen in recA nicht nur hinsichtlich ihrer genetischen Stabilität sondern auch ihrer Sicherheit verbessert werden. Dies gilt, wie unten weiter ausgeführt, insbesondere für Stämme, die in der biotechnologischen Produktion tatsächlich eingesetzt werden, wie beispielsweise *Bacillus licheniformis*.

Wie unten ebenfalls detaillierter ausgeführt geschieht dies vorzugsweise im Zusammenhang mit einer und besonders bevorzugt keinen weiteren sicherheitsrelevanten Deletionen. Ebenso bevorzugt geschieht dies in möglichst nahe verwandten Stämmen. Hierbei ist jedoch von *B. megaterium* abzusehen, zum einen weil hierfür, wie einleitend

beschrieben, bereits die Spezies-eigenen recA-Gene beziehungsweise die hierin deletierten Mutanten zur Verfügung stehen und zum anderen weil diese Spezies, die sich insbesondere durch ihre großen Zellen und damit verbundene mikrobiologische Eigenheiten auszeichnet, im allgemeinen nicht für die großtechnische Fermentation genutzt wird.

Gegenstände der vorliegenden Erfindung liegen somit in dem Faktor RecA (SEQ ID NO. 2) und dem zugehörigen Gen recA (SEQ ID NO. 1) aus *B. licheniformis* DSM 13 beziehungsweise nahen Verwandten hierzu. Ebenso stellt die Verwendung eines recA-Gens zu dessen funktioneller Inaktivierung in einem grampositiven Bakterium einen Erfindungsgegenstand dar, vorzugsweise in Kombination mit der funktionellen Inaktivierung eines in der Phase IV der Sporulation grampositiver Mikroorganismen aktiven Gens, vorzugsweise *spolV*, *yqfD* beziehungsweise Homologen hierzu. Dies geschieht vorteilhafterweise mithilfe der in der vorliegenden Anmeldung zusätzlich beschriebenen Gene *spolV* und *yqfD*. Einen entsprechenden Gegenstand der vorliegenden Erfindung stellen die hierdurch erhaltenen grampositiven Mikroorganismen dar; ebenso die mit diesen Organismen durchgeführten Fermentationen, insbesondere zur Herstellung von Wertstoffen. Des weiteren steht mit der vorliegenden Anmeldung ein RecA-Protein zur Verfügung, das in molekularbiologischen Ansätzen oder zur Modulation der molekularbiologischen Aktivitäten von Zellen verwendet werden kann, insbesondere im Zusammenhang mit DNA-Polymerisations- oder Rekombinationsvorgängen.

Zum ersten Erfindungsgegenstand gehört jeder oben definierte Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz zunehmend bevorzugt mindestens zu 96,5%, 97%, 97,5%, 98%, 98,5%, 99%, 99,5% und ganz besonders bevorzugt zu 100% identisch ist.

Denn wie erläutert ist mit zunehmender Ähnlichkeit eine zunehmende Übereinstimmung der Funktionen und damit eine Austauschbarkeit der Faktoren zu erwarten.

Vorzugsweise handelt es sich dabei um einen Faktor RecA, der von einer Nukleinsäure codiert ist, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.

In bevorzugten Ausführungsformen sind das Faktoren, die von einer Nukleinsäure codiert sind, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.

Denn über die Nukleinsäuren stehen die betreffenden Faktoren beziehungsweise Gene für die Transformation in andere, vorzugsweise verwandte Spezies oder für Modifikationen zur Verfügung. Hierzu gehören, wie unten detaillierter erläutert, insbesondere Mutationen der betreffenden Gene. Mit einem zunehmenden Maß an Identität zur angegebenen Sequenz sollte der Erfolg bei solchen Spezies umso größer sein, die zu B. licheniformis zunehmend verwandt sind, insbesondere bei der für die biotechnologische Produktion besonders wichtigen Spezies B. licheniformis selbst.

Wie erläutert dienen der Verwirklichung der vorliegenden Erfindung vor allem die für einen Faktor RecA codierenden Nukleinsäuren, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.

Dies gilt um so mehr für derartige Nukleinsäuren, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist. Denn diese können über entsprechende Konstruktionen zur Transformation und/oder zur Mutagenese verwendet werden, wobei eine zunehmende Ähnlichkeit den erwünschten Erfolg umso wahrscheinlicher werden lässt.

Ganz besonders bevorzugt handelt es sich dabei um eine derartige Nukleinsäure, die für einen zuvor beschriebenen Faktor RecA codiert. Dies gilt beispielsweise für Strategien, bei denen ein funktioneller Faktor RecA hergestellt werden soll, etwa für die unten ausgeführten molekularbiologischen Versuchsansätze, oder für die Erzielung einer maximalen Übereinstimmung mit dem jeweils endogenen tatsächlich für RecA codierenden Gen, welches modifiziert und/oder ausgeschaltet werden soll. Denn in vielen Fällen reicht eine Mutation in einer einzigen Position, um etwa über eine nonsense-Mutation das Gen beziehungsweise den Faktor in seiner natürlichen Funktion auszuschalten.

Einen eigenen Erfindungsgegenstand stellt die Verwendung einer für einen Faktor RecA codierenden Nukleinsäure zur funktionellen Inaktivierung des Gens recA in einem grampositiven Bakterium dar, welches nicht *Bacillus megaterium* ist.

Denn zum einen gibt es für *B. megaterium* bereits die einleitend erwähnten Studien, in denen vorgeschlagen wird, gleichzeitig mit mehreren anderen Mutationen auch recA zu deletieren, um zu Sicherheitsstämmen zu gelangen. Zum zweiten stellen andere grampositive Bakterien wie beispielsweise solche der Gattungen *Bacillus*, *Staphylococcus*, *Corynebakterium* und *Clostridium* im Stand der Technik wichtigere Wirtsorganismen für die biotechnologische Produktion von Wertstoffen (siehe unten) dar.

Unter der funktionellen Inaktivierung ist im Sinne der vorliegenden Anmeldung jede Art von Modifikation oder Mutation zu verstehen, wonach die Funktion eines RecA als Einzelstrang-DNA-bindenden Faktors unterbunden wird. Dazu gehört die Ausführungsform, daß ein praktisch vollständiges, aber inaktives Protein gebildet wird, daß inaktive Teile eines RecA in der Zelle vorliegen, bis hin zu den Möglichkeiten, daß das Gen recA nicht mehr translatiert wird oder sogar vollständig deletiert ist. Somit besteht die eingangs diskutierte „Verwendung“ dieses Faktors oder dieses Gens dieser Ausführungsform nach darin, daß er beziehungsweise es von der betreffenden Zelle eben nicht mehr auf seine natürliche Weise zur Wirkung kommt. Dies wird diesem Erfindungsgegenstand zufolge auf genetischer Ebene dadurch erreicht, daß das betreffende Gen ausgeschaltet wird.

Nach einer Ausführungsform dieser Verwendung wird eine für ein nichtaktives Protein codierende Nukleinsäure mit einer Punktmutation eingesetzt.

Derartige Nukleinsäuren können über an sich bekannte Verfahren zur Punktmutagenese erzeugt werden. Solche sind beispielsweise in einschlägigen Handbüchern wie dem von Fritsch, Sambrook und Maniatis, „Molecular cloning: a laboratory manual“, Cold Spring Harbour Laboratory Press, New York, 1989, dargestellt. Zudem stehen hierfür inzwischen zahlreiche kommerzielle Baukästen zur Verfügung, etwa das QuickChange®-Kit der Firma Stratagene, La Jolla, USA. Das Prinzip besteht darin, daß Oligonukleotide mit einzelnen Austauschen (Mismatch-Primer) synthetisiert und mit dem einzelsträngig vorgelegten Gen hybridisiert werden; anschließende DNA-Polymerisation ergibt dann entsprechende Punktmutanten. Hierfür können die jeweiligen Spezies-eigenen recA-Sequenzen verwendet werden. Aufgrund der hohen Homologien ist es möglich und erfindungsgemäß

besonders vorteilhaft, diese Reaktion anhand der mit SEQ ID NO. 1 zur Verfügung gestellten Sequenz oder etwa den anderen aus Figur 2 hervorgehenden Sequenzen verwandter Spezies durchzuführen. Diese Sequenzen können auch dazu dienen, entsprechende Mismatch-Primer für verwandte Spezies zu entwerfen, insbesondere anhand der im Alignment der Figur 2 identifizierbaren konservierten Bereiche.

Nach einer Ausführungsform dieser Verwendung wird eine Nukleinsäure mit einer Deletions- oder Insertionsmutation eingesetzt, vorzugsweise umfassend die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassenden Randsequenzen des für das Protein codierenden Bereichs.

Auch diese Verfahren sind dem Fachmann an sich vertraut. Somit ist es möglich, die Bildung eines Faktors RecA durch die Wirtszelle dadurch zu verhindern, daß ein Teil des Gens auf einem entsprechenden Transformationsvektor über Restriktionsendonukleasen herausgeschnitten und der Vektor anschließend in den interessierenden Wirt transformiert wird, wo über die – bis dahin noch mögliche – homologe Rekombination das aktive Gen gegen die inaktive Kopie ausgetauscht wird. In der Ausführungsform der Insertionsmutation kann lediglich das intakte Gen unterbrechend oder anstelle eines recA-Genteils ein anderes Gen, beispielsweise ein Selektionsmarker eingefügt werden. Hierüber ist das Mutationsereignis in an sich bekannter Weise phänotypisch überprüfbar.

Um diese jeweils notwendigen Rekombinationsergebnisse zwischen dem in die Zelle eingeführten defekten Gen und der beispielsweise auf dem Chromosom endogen vorhandenen intakten Genkopie zu ermöglichen, ist nach dem derzeitigen Wissensstand eine Übereinstimmung in jeweils mindestens 70 bis 150 zusammenhängenden Nukleinsäurepositionen, jeweils in den beiden Randsequenzen zu dem nichtübereinstimmenden Teil nötig, wobei es auf den dazwischenliegenden Teil nicht ankommt. Dementsprechend sind solche Ausführungsformen bevorzugt, die lediglich zwei flankierende Regionen mit mindestens diesen Größen umfassen.

Nach einer alternativen Ausführungsform dieser Verwendung werden Nukleinsäuren mit insgesamt zwei Nukleinsäureabschnitten eingesetzt, die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassen und damit den für das Protein codierenden Bereich zumindest teilweise, vorzugsweise vollständig flankieren.

Denn allein um den Austausch der beiden Genkopien über homologe Rekombination zu ermöglichen, braucht es sich dabei nicht zwangsläufig um proteincodierende Abschnitte zu handeln. Vielmehr eignen sich hierfür auch die Randbereiche der betreffenden Gene, welche natürlicherweise eine andere Funktion (Promotor, Terminator, Enhancer etc.) ausüben oder lediglich nichtfunktionelle intergenische Abschnitte darstellen. So kann die funktionelle Inaktivierung beispielsweise auch in der Deletion des Promotors bestehen, wofür es bei einer Deletionsmutation dieser Ausführungsform notwendig ist, auf flankierende, nichtcodierende Abschnitte zurückzugreifen. Je nach Einzelfall kann es auch sinnvoll sein, für die flankierenden Regionen solche Abschnitte auszuwählen, die zum Teil in den proteincodierenden Bereich hineinreichen und zum Teil außerhalb liegen.

Derartige, wenigstens zum Teil nichtcodierende Bereiche können für die Gene recA aus *B. amyloliquefaciens* und für recA und recE aus *B. subtilis* den oben angegebenen Datenbankeinträgen entnommen werden. Für andere Stämme, beispielsweise auch für recA aus *B. licheniformis* ist es möglich; die betreffenden nichtcodierenden Bereiche über PCR-basierte Verfahren aus einer Präparation der genomischen DNA zu erschließen. Diese Verfahren (beispielsweise anchored PCR) sind im Stand der Technik etabliert. Als Ausgangspunkte hierfür dienen die bekannten Genabschnitte. Der vorliegenden Erfindung zufolge können die hierfür benötigten Primer anhand der SEQ ID NO. 1 auch für andere Spezies grampositiver Bakterien und hierunter insbesondere für solche der Gattung *Bacillus* entworfen werden.

Bei einer entsprechenden Verwendung handelt es sich demzufolge um eine der zuvor beschriebenen erfindungsgemäßen Nukleinsäuren beziehungsweise um die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren. Die verschiedenen hierunter fallenden Ausführungsformen sind entsprechend bevorzugt.

In bevorzugten Ausführungsformen dieser Verwendungen handelt es sich bei dem grampositiven Bakterium vorzugsweise um eines der Gattungen *Clostridium* oder *Bacillus* und um eines, das natürlicherweise zur Sporulation befähigt ist, bei dem gleichzeitig mit recA ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.

Denn viele grampositiven Bakterien sind wie einleitend beschrieben in der Lage, unter entsprechend ungünstigen Umweltbedingungen den Vorgang der Sporulation einzuleiten. Dies kann erfindungsgemäß insofern für Sicherheitsaspekte ausgenutzt werden, als in

Kombination mit der oben dargestellten funktionellen Inaktivierung von *recA* ein Sporulationsgen aus der vergleichsweise späten Phase IV der Sporulation ebenfalls funktionell inaktiviert wird. Damit stehen der formulierten Aufgabe entsprechend zwei gleichzeitig und erfindungsgemäß miteinander kombinierbare Systeme für die Herstellung sicherer GVO zur Verfügung. Die Kombination beider Systeme war bislang noch nicht bekannt, insbesondere nicht zu diesem Zweck.

Besonders erfolgreich konnte die Inaktivierung von Sporulationsgenen an Spezies der Gattungen *Clostridium* und *Bacillus* realisiert werden, weshalb die hierdurch gekennzeichneten Ausführungsformen entsprechend bevorzugt sind.

Insbesondere ist es überraschend, daß die Verhinderung der Sporulation erst in einem so späten Stadium für diesen Zweck erfolgreich ist. Zwar werden in *spoIV*-Mutanten von *B. licheniformis* unter entsprechenden Bedingungen noch Sporen (die sogenannten „Phase-Grau-Sporen“) gebildet, doch sind diese steril und nicht mehr in der Lage auszukeimen. Insofern trägt diese Mutation dem Sicherheitsaspekt Rechnung. Bisher war eine Unterbindung der Sporulation eher zu einem früheren Zeitpunkt favorisiert worden. Die Inaktivierung in Phase IV sorgt zusätzlich jedoch dafür, daß die in den früheren Sporulations-Phasen aktiven Faktoren auch weiterhin in den Mutanten gebildet werden. Ohne an diese Theorie gebunden sein zu wollen, kann man vermuten, daß zumindest einige dieser Faktoren von den Zellen auch für die normalen während der Fermentation ablaufenden Stoffwechselvorgänge benötigt werden. Bei dem Ausschalten zu einem früheren Zeitpunkt stünden sie nicht mehr zur Verfügung. Umgekehrt besteht die vorteilhafte Wirkung der Inaktivierung der Sporulation in Phase IV darin, daß der hierdurch stattfindende Eingriff in die Physiologie der Zellen nicht so gravierend ist und die Fermentation an sich weniger beeinträchtigt wird als bei einem früheren Ausschalten dieser Gene.

Bevorzugt ist eine derartige Verwendung, bei der es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIV*A, *spoIV*B, *spoIV*C, *spoIV*CB, *spoIV*FA, *spoIV*FB oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

All diese Gene sind an sich bekannt und für diese Phase der Sporulation beschrieben. Das *B. subtilis*-Gen *spolVA* codiert für das Phase-IV-Sporulationsprotein A, das in den Datenbanken Swiss-Prot (Geneva Bioinformatics (GeneBio) S.A., Genf, Schweiz; <http://www.genebio.com/sprot.html>) und NCBI (siehe oben) unter der Nummer P35149 hinterlegt ist. Es spielt für die Bildung einer intakten Sporenhülle und deren Zusammenbau eine Rolle. Die Aminosäuresequenz des zugehörigen Faktors *SpoIVVA* wird in SEQ ID NO. 8 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist des Institute Pasteur, Paris, Frankreich (<http://genolist.pasteur.fr/SubtiList/genome.cgi>) unter der Nummer BG10275 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 7 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 7 angegebene Nukleotidsequenz von 1 bis 1876 erfindungsgemäß als Gen *spolVA* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1679; das erste Codon, das heißt die Positionen 201 bis 203 werden *in vivo* nicht als Leucin sondern als Methionin translatiert.

Das *B. subtilis*-Gen *spolVB* codiert für das Phase-IV-Sporulationsprotein B, das in den Datenbanken Swiss-Prot und NCBI unter der Nummer P17896 hinterlegt ist. Es ist für den Sigma-Faktor-K-abhängigen Übergangspunkt während der Sporulation beziehungsweise dessen Aktivierung in der Mutterzelle von Bedeutung. Es spielt für die interkompartimentelle Signalübertragung eine Rolle, wahrscheinlich über den hydrophoben N-Terminus. Die Aminosäuresequenz des Faktors *SpoIVB* wird in SEQ ID NO. 10 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10311 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 9 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 9 angegebene

Nukleotidsequenz von 1 bis 1675 erfindungsgemäß als Gen *spolVB* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1478.

Das *B. subtilis*-Gen *spolVCA* codiert für eine putative ortsspezifische DNA-Rekombinase, die in den Datenbanken Swiss-Prot und NCBI unter der Nummer P17867 hinterlegt ist. Sie spielt wahrscheinlich eine Rolle, um die Gene *spolIIC* und *spolVCB* zu rekombinieren, woraus der Sigmafaktor K hervorgeht. Die Aminosäuresequenz dieser Rekombinase wird in SEQ ID NO. 12 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10458 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 11 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 11 angegebene Nukleotidsequenz von 1 bis 1900 erfindungsgemäß als Gen *spolVCA* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1703; das erste Codon, das heißt die Positionen 201 bis 203 werden *in vivo* nicht als Valin sondern als Methionin translatiert.

Das *B. subtilis*-Gen *spolVCB* codiert für den RNA-Polymerase-Sigmafaktor-K-Precursor, der in den Datenbanken Swiss-Prot und NCBI unter der Nummer P12254 hinterlegt ist. Der Rest dieses Faktors wird von dem Gen *spolIIC* codiert, welches auf dem Chromosom ca. 10 kb entfernt ist, wobei der dazwischenliegende Bereich als SKIN bezeichnet wird. Durch Excision dieses Fragments in der unmittelbar vorangehenden Sporulationsphase wird der aktive Sigma-Faktor K erhalten, welcher seinerseits als Transkriptionsfaktor wirkt. Die Aminosäuresequenz des Teilstücks *SpolVCB* wird in SEQ ID NO. 14 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10459 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 13 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 13 angegebene Nukleotidsequenz von

1 bis 868 erfindungsgemäß als Gen spolVCB bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 671.

Das B. subtilis-Gen spolVFA codiert für das Phase IV-Sporulationsprotein FA. Dieser Faktor, der vermutlich in der Lage ist, mit SpolVFB (siehe unten) ein Heterodimer zu bilden, erfüllt wahrscheinlich die Aufgabe, diesen Faktor zu stabilisieren aber dadurch gleichzeitig auch zu inhibieren. Deshalb wird SpolVFA auch schon zu einem früheren Zeitpunkt, vermutlich in Phase II gebildet. Die Aminosäuresequenz von SpolVFA ist in den Datenbanken Swiss-Prot und NCBI unter der Nummer P26936 hinterlegt. Sie wird in SEQ ID NO. 16 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10331 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 15 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 15 angegebene Nukleotidsequenz von 1 bis 1192 erfindungsgemäß als Gen spolVFA bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 995.

Das B. subtilis-Gen spolVFB codiert für das Phase-IV-Sporulationsprotein FB. Dabei handelt es sich um eine membranassoziierte Metalloprotease, die vermutlich für die Prozessierung von Pro-Sigma K zu Sigma K zuständig ist; sie wird ebenfalls bereits in Phase II der Sporulation gebildet. Die Aminosäuresequenz von SpolVFB ist in den Datenbanken Swiss-Prot und NCBI unter der Nummer P26937 hinterlegt. Sie wird in SEQ ID NO. 18 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10332 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 17 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 17 angegebene Nukleotidsequenz von 1 bis 1264 erfindungsgemäß als Gen spolVFB bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1067; das erste Codon, das

heißt die Positionen 201 bis 203 werden *in vivo* nicht als Leucin sondern als Methionin translatiert.

Auch die beiden bevorzugten Gene gehen an sich aus dem Stand der Technik hervor. Die DNA- und Aminosäuresequenzen von *spoIV* aus *B. licheniformis* sind in der Datenbank NCBI unter der Nummer AJ616332 hinterlegt. Dieser Faktor wird in der Diplomarbeit „Arbeiten zur Herstellung einer sporulationsnegativen Mutante von *Bacillus licheniformis*“ von M. Gröne (2002), im Fachbereich Biologie der Westfälischen Wilhelms-Universität Münster, Deutschland als ein für die Sporulation von *B. licheniformis* essentieller Faktor beschrieben. Die zugehörigen, zum Teil auch regulatorische Bereiche umfassenden Sequenzen sind in der vorliegenden Anmeldung unter SEQ ID NO. 3 und 4 angegeben. Zu diesen Sequenzen ist zu bemerken, daß erfindungsgemäß der Bereich der Nukleotide 1 bis 1792 als Gen *spoIV* bezeichnet wird, wobei der eigentliche *SpoIV*-codierende Abschnitt die Positionen 140 bis 1336 umfaßt; die Randsequenzen mögen wiederum andere genetische Elemente wie Regulationselemente oder Teile anderer Gene enthalten. Hierunter wird das erste Codon GTG der Positionen 140 bis 142 *in vivo* nicht als Valin sondern als Methionin translatiert.

In dieser Arbeit wird auch auf den Faktor beziehungsweise das Gen *yqfD* aus *B. subtilis* hingewiesen, welches mit einer Homologie von 68% Identität auf Aminosäureebene als das nächstähnliche bis dato bekannte Protein angesehen wird. Dieser Faktor ist in der Datenbank Swiss-Prot unter der Nummer P54469 angegeben; sowohl die Aminosäuresequenz als auch die DNA-Sequenz mit beiden ca. 200 bp flankierenden Bereichen gehen aus der Datenbank Subtilist unter der Nummer BG11654 hervor. Der dortige Eintrag vermerkt, es sei zwar ein unbekanntes Protein, aber aufgrund der bestehenden Sequenzhomologien könne es als Ähnliches zum Phase-IV-Sporulationsprotein angesehen werden. Die zugehörigen Sequenzen können SEQ ID NO. 5 und 6 der vorliegenden Anmeldung entnommen werden. Zu diesen Sequenzen ist ergänzend zu bemerken, daß ungeachtet der ebenfalls angegebenen und möglicherweise anderen genetischen Elemente enthaltenden Randssequenzen erfindungsgemäß der Bereich der Nukleotide 1 bis 1594 als Gen *yqfD* bezeichnet wird, wobei der eigentliche proteincodierende Abschnitt die Positionen 201 bis 1397 umfaßt. Hierunter wird das erste Codon GTG der Positionen 201 bis 203 wie bei *spoIV* aus *B. licheniformis* *in vivo* nicht als Valin sondern als Methionin translatiert.

Es ist zu erwarten, daß alle anderen grampositiven, natürlicherweise zur Sporulation befähigten Mikroorganismen über Homologe zu den genannten sieben Genen und davon abgeleitete Faktoren vergleichbarer Funktionen verfügen. Sie dürften in an sich bekannter Weise unter Hybridisierung mit den in im Sequenzprotokoll angegebenen Nukleinsäuren ohne weiteres identifizierbar sein, insbesondere mithilfe der beiden homologen Sequenzen SEQ ID NO. 3 beziehungsweise 5, womit eine gewisse Varianz über Spezies-Grenzen hinweg ermöglicht wird.

Eines dieser Gene, vorzugsweise yqfD / spoIV beziehungsweise dessen Homologes wird im Produktionsstamm erfindungsgemäß gleichzeitig mit recA inaktiviert, um hieraus entsprechende Sicherheitsstämme zu erhalten. Vorteilhafterweise werden hierfür entsprechend den oben gemachten Ausführungen zur Deletionsmutagenese die in SEQ ID NO. 3, 5, 7, 9, 11, 13, 15 oder 17 angegebenen Nukleinsäuren selbst zur Inaktivierung verwendet. Damit ist es nicht einmal nötig, die betreffenden homologen Gene aus den für die Produktion verwendeten Spezies selbst zu identifizieren. Hierbei ist zu erwarten, daß diese Deletionen umso erfolgreicher sind, je enger die betreffenden Spezies mit *B. subtilis* beziehungsweise *B. licheniformis* verwandt sind. Denn hiermit sollte eine zunehmende Homologie der betreffenden Gene verbunden sein. Aus diesem Grund sind im Sequenzprotokoll jeweils auch die ca. 200 bp umfassenden Randsequenzen angegeben, denn damit können entsprechend den für recA gemachten Ausführungen Konstrukte gebildet werden, die die für ein Crossing-over nötigen mindestens 70 bis 150 Positionen umfassenden Bereiche in vollständig flankierenden Bereichen enthalten.

Erfindungsgemäß ist es möglich, zusammen mit recA mehrere der genannten Phase IV-Gene zu inaktivieren und dadurch Sicherheitsstämme zu erhalten, die neben der Unfähigkeit zur RecA-vermittelten DNA-Rekombination nicht in der Lage sind, reife Sporen zu bilden. Erfindungsgemäß reicht es hierfür jedoch aus, neben recA lediglich eines dieser Gene zu inaktivieren, weshalb in einer bevorzugten derartigen Verwendung genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.

Hierdurch werden Sicherheitsstämme für die biotechnologische Produktion erhalten. Diese sind außerhalb der optimalen Fermentationsbedingungen weniger gut überlebensfähig, insbesondere unter Umweltbedingungen, die schlechte Nährstoffversorgung und DNA-schädigende Einflüsse, etwa durch UV-Strahlung oder aggressive chemische Verbindungen, umfassen. Die genannte erste Gruppe von Umwelteinflüssen würde bei

natürlicherweise zur Sporulation befähigten grampositiven Bakterien den Übergang in die Dauerform der Sporen induzieren; die zweite Gruppe von Einflüssen kann von Mikroorganismen natürlicherweise über RecA-vermittelte DNA-Reparatur- und Rekombinationsprozesse ausgeglichen werden. Wenn die Zellen zu beidem nicht mehr oder nur noch stark eingeschränkt in der Lage sind, sind sie erfindungsgemäß als Sicherheitsstämme geeignet.

Ferner ist es möglich, zusammen mit recA eines oder mehrere der im Stand der Technik bekannten Gene zu inaktivieren und dadurch Sicherheitsstämme zu erhalten, die neben der Unfähigkeit zur RecA-vermittelten DNA-Rekombination und zur Bildung reifer Sporen auch über diese zusätzlichen Eigenschaften gekennzeichnet sind. Dazu gehören neben „aktiven“, die Lebensfähigkeit unterbindenden und entsprechend stringent zu regulierenden Systemen auch jene, die im einleitend dargestellten Stand der Technik als „passive“ Systeme zur Erzeugung von GVO bezeichnet worden sind. dazu gehören insbesondere inaktivierende Mutationen in einem oder mehreren der folgenden Gene: epr, rp-I, rp-II, isp-1, apr, npr, spoOA, bpr, rsp, mpr, vpr, spo0A, spoll:D, spollAC, spo2, spo3, sigE, sigF, spollE, spollSB, sigG, spollVCB, spollIC, nprM und das Gen für die Isopropylmalat-Dehydrogenase (leuB). Diese Genbezeichnungen sind dem in der Einleitung zur vorliegenden Anmeldung dargestellten Stand der Technik entnommen. Somit sind an dieser Stelle mit diesen Abkürzungen jeweils die in den einleitend genannten Anmeldungen und Publikationen beschriebenen Bedeutungen gemeint. Sollten für dieselben Gene beziehungsweise Gengruppen, codierend für dieselben oder homologe Proteine, im Stand der Technik weitere Namen etabliert sein, insbesondere für die Homologen in anderen Bakterienspezies als denen, die den erwähnten Arbeiten zugrundegelegen haben, so gilt das hier Gesagte entsprechend.

Erfindungsgemäß ist es jedoch nicht unbedingt notwendig, neben recA und gegebenenfalls zusätzlich einem Sporulationsphasen IV-Gen ein weiteres Gen zu inaktivieren, so daß vorzugsweise auf diese weiteren Mutationen weitgehend verzichtet wird. Hiermit ist der in der Aufgabe zur vorliegenden Anmeldung geforderte Vorteil verbunden, möglichst wenige Sicherheitssysteme parallel in derselben Zelle zu etablieren. Hiermit wird der Arbeitsaufwand geringer gehalten, als wenn man, wie in den genannten Arbeiten zu B. megaterium vorgeschlagen, vier verschiedene Deletionen vornehmen müßte. Das ist insbesondere dann relevant, wenn die betreffenden Zellen zuerst - solange sie noch zur Rekombination in der Lage sind - mit den für die Produktion relevanten

Transgenen versehen werden und dann erst in Sicherheitsstämme, insbesondere in einen recA-Minus-Phänotyp überführt werden. Nur für sehr kritische Fälle, etwa hochpathogene Stämme, sind derartige weitere Mutationen angezeigt.

Aufgrund der mit der vorliegenden Anmeldung zur Verfügung gestellten Sequenzen werden Sporulationsdefekte in den Genen *spoIV*A, *spoIV*B, *spoIV*C, *spoIV*C, *spoIV*F, *spoIV*F oder *yqfD* in der Nomenklatur von *B. subtilis* beziehungsweise im Fall von *Bacillus licheniformis* in dem Gen *spoIV* beziehungsweise den je nach Wirtszelle vorhandenen hierzu homologen Genen erzeugt.

Diese Homologie läßt sich in erster Näherung über einen Sequenzvergleich erschließen. Zur Kontrolle kann im vorliegenden, für die biotechnologische Produktion vorgesehenen Mikroorganismus-Stamm das fragliche Gen inaktiviert und über eine Wiederherstellung des Phänotyps (Rescue) die funktionelle Übereinstimmung der betreffenden Gene überprüft werden. Überführt die parallele Bereitstellung einer erfindungsrelevanten *spoIV*A, *spoIV*B, *spoIV*C, *spoIV*C, *spoIV*F, *spoIV*F, *yqfD*- oder *spoIV*-Kopie die betreffende Knock-out-Mutante wieder in einen Sporulations-positiven Phänotyp, so wäre damit der Nachweis erbracht, daß auch eine funktionelle Austauschbarkeit der betrachteten Gene besteht. Unter homologen Genen zu den genannten Phase-IV-Sporulationsgenen werden erfindungsgemäß also insbesondere solche verstanden, die einem derartigen „Rescue“ zugänglich sind. Wenn das möglich ist, handelt es sich um ein bevorzugt verwendetes Sporulationsgen. Diese Kontrolle ist insbesondere deshalb mit zumutbarem Arbeitsaufwand möglich, weil zum einen erfindungsgemäß eben eine solche funktionell inaktive Mutante erzeugt werden soll und zum anderen über das Sequenzprotokoll zur vorliegenden Anmeldung die betreffenden Sequenzen aus *B. subtilis* und die besonders bevorzugte davon zusätzlich aus *B. licheniformis* zur Verfügung gestellt werden, über die ein derartiger Rescue vorgenommen werden kann.

In bevorzugten Ausführungsformen erfolgt die erfindungsgemäße, bisher beschriebene Verwendung zur funktionellen Inaktivierung der Gene *spoIV*A, *spoIV*B, *spoIV*C, *spoIV*C, *spoIV*F, *spoIV*F, *yqfD* oder *spoIV* beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

Denn wie oben beschrieben können über diese konkreten Sequenzen, insbesondere für *B. licheniformis* und *B. subtilis* und nahe mit diesen verwandte Spezies entsprechende molekularbiologische Konstrukte hergestellt werden. Hierfür stehen alle oben zu *recA* ausgeführten Möglichkeiten zur Verfügung und sind entsprechend bevorzugt.

Einen eigenen Gegenstand der vorliegenden Erfindung stellen die mit den beschriebenen Verfahren erhaltenen Mikroorganismen dar. In seiner allgemeinsten Formulierung handelt es sich dabei also um ein grampositives Bakterium, welches nicht *Bacillus megaterium* ist, bei dem das Gen *recA* funktionell inaktiviert ist.

Denn wie bereits gesagt sind grampositive Bakterien, beispielsweise aufgrund ihrer Fähigkeit zur Sekretion von Wertstoffen und/oder ihrer vergleichsweise leichten Fertigierbarkeit die für die Biotechnologie wichtigsten Mikroorganismen. Hierunter werden für die verschiedenen Einsatzgebiete verschiedene Spezies bevorzugt, so werden niedermolekulare Verbindungen wie etwa Aminosäuren in besonders großem Ausmaß mithilfe von *Corynebacterien* produziert; *Bacillus* und hierunter insbesondere *B. licheniformis* wird für die Produktion von extrazellulären Proteinen besonders geschätzt. Sie alle sind erfindungsgemäß einer funktionellen Inaktivierung von *RecA* zugänglich.

Diese erfindungsgemäßen Bakterien zeichnen sich durch die beschriebenen Rekombinationsdefekte aus und weisen deshalb unter natürlichen Bedingungen, insbesondere in Konkurrenz mit anderen Mikroorganismen Nachteile hinsichtlich ihrer Überlebensfähigkeit auf und eignen sich somit als Sicherheitsstämme für die biotechnologische Produktion. Hierbei handelt es sich aus den oben ausgeführten Gründen nicht um *Bacillus megaterium*.

Den bisherigen Ausführungen entsprechend sind darunter solche grampositiven Bakterien bevorzugt, bei denen die funktionelle Inaktivierung über Punktmutagenese, teilweise Deletion oder Insertion oder vollständige Deletion des für das vollständige Protein codierenden Bereichs erfolgt ist.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien bevorzugt, bei denen die funktionelle Inaktivierung über eine erfindungsgemäße für *RecA* codierende Nukleinsäure beziehungsweise über die zumindest teilweise

nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren erfolgt ist. Hierbei sind die Nukleinsäuren den oben beschriebenen Homologiewerten entsprechend bevorzugt. Mikroorganismen, bei denen die funktionelle Inaktivierung mithilfe der in SEQ ID NO. 1 angegebenen Nukleinsäure beziehungsweise Abschnitten davon erfolgt ist, stellen in dieser Hinsicht die am meisten bevorzugten Mikroorganismen dar. Entsprechend dem oben gesagten sind im Falle einer Mutagenese über Crossing over vorzugsweise Randsequenzen von jeweils mindestens 70 bis 150 bp verwendet worden, was über eine Sequenzierung der betreffenden chromosomalen Abschnitte überprüft werden kann.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien und hierunter vorzugsweise solche der Gattungen Clostridium oder Bacillus, bevorzugt, die natürlicherweise zur Sporulation befähigt sind und bei denen gleichzeitig mit recA ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien bevorzugt, bei denen es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB* oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

In besonderen Fällen, etwa beim Einsatz hochpathogener Stämme können auch mehrere der genannten Sporulationsgene oder eines oder mehrere der im Stand der Technik beschriebenen Gene oder Gengruppen *spoIV/yqfD/Homolog*, *epr*, *rp-I*, *rp-II*, *isp-1*, *apr*, *npr*, *spoOA*, *bpr*, *rsp*, *mpr*, *vpr*, *spoOA*, *spoII:D*, *spoIIAC*, *spo2*, *spo3*, *sigE*, *sigF*, *spoIIIE*, *spoIIISB*, *sigG*, *spoIVCB*, *spoIIIC*, *nprM* und/oder das Gen für die Isopropylmalat-Dehydrogenase (*leuB*) funktionell inaktiviert werden. Unter diesen Abkürzungen sind jeweils die in den einleitend genannten Anmeldungen und Publikationen beschriebenen Bedeutungen zu verstehen, wobei eventuelle Synonyme entsprechend eingeschlossen werden.

Den bisherigen Ausführungen entsprechend handelt es sich dabei jedoch bevorzugt jeweils um ein solches grampositives Bakterium, bei dem – neben der RecA-Inaktivierung – genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.

Entsprechend dem oben Gesagten ist weiterhin jeweils ein solches derartiges grampositives Bakterium bevorzugt, bei dem die funktionelle Inaktivierung der Gene *spolVA*, *spolVB*, *spolVCA*, *spolVCB*, *spolVFA*, *spolVFB*, *yqfD* oder *spolV* beziehungsweise den hierzu jeweils homologen Genen mithilfe einer der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt ist, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

Dies läßt sich über Präparationen der betreffenden DNA, beispielsweise der chromosomalen DNA eines erfindungsgemäßen Stamms und Restriktionsanalyse oder PCR überprüfen. Als Primer hierfür können die jeweiligen flankierenden Sequenzen verwendet werden, wobei die Größe des PCR-Produkts Aufschluß über das Vorhandensein und gegebenenfalls die Größe von Inserts gibt.

Entsprechend den bisherigen Ausführungen sind unter den erfindungsgemäßen grampositiven Bakterien besonders solche bevorzugt, bei denen es sich um Vertreter der Gattungen *Clostridium* oder *Bacillus* handelt, insbesondere um solche der Spezies *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. stearothermophilus*, *B. globigii*, *B. clausii* oder *B. lentus*, und ganz besonders um Stämme von *B. licheniformis*.

Einen eigenen Erfindungsgegenstand stellen die Verfahren zur Fermentation eines erfindungsgemäßen grampositiven Bakteriums dar.

Denn diese Verfahren zeichnen sich dadurch aus, daß *RecA*, vorzugsweise im Zusammenspiel mit den dargestellten bevorzugten Ausführungsformen nicht aktiv ist und der betreffende Stamm im Falle einer versehentlichen Freisetzung in die Umgebung der Anlage ein deutlich minimiertes Sicherheitsrisiko darstellt. Für Verfahren zur Fermentation werden entsprechende Sicherheitsanforderungen gestellt, so daß sie gegebenenfalls nur dann durchführbar sind, wenn sie diese Anforderungen erfüllen.

Vorzugsweise handelt es sich dabei um Verfahren zur Herstellung eines Wertstoffs, insbesondere einer niedermolekularen Verbindung oder eines Proteins.

Zwar ist es vorteilhaft, wenn entsprechende Stämme auch im Labormaßstab eingesetzt werden. Doch sind hier die übrigen Rahmenbedingungen im allgemeinen leichter zu erfüllen. Außerdem besteht das Haupeinsatzgebiet der Fermentation von Mikroorganismen in der biotechnologischen Herstellung von Wertstoffen.

Der Bedeutung dieser Wertstoffe entsprechend sind hierunter solche Verfahren bevorzugt, bei denen es sich bei der niedermolekularen Verbindung um einen Naturstoff, einen Nahrungsmittelergänzungsstoff oder um eine pharmazeutisch relevante Verbindung handelt.

Hierunter sind beispielsweise Aminosäuren oder Vitamine zu nennen, die besonders als Nahrungsmittelergänzungsstoffe Verwendung finden. Bei pharmazeutisch relevanten Verbindungen kann es sich um Vor- oder Zwischenstufen zu Medikamenten oder sogar um diese selbst handeln. In all diesen Fällen spricht man auch von Biotransformation, wonach die Stoffwechseleigenschaften der Mikroorganismen ausgenutzt werden, um die ansonsten aufwendige chemische Synthese ganz oder zumindest in einzelnen Schritten zu ersetzen.

In nicht minder bevorzugten Verfahren handelt es sich bei dem Protein um ein Enzym, insbesondere eines aus der Gruppe der α -Amylasen, Proteasen, Cellulasen, Lipasen, Oxidoreduktasen, Peroxidasen, Laccasen, Oxidasen und Hemicellulasen.

Industrielle Enzyme, die mit derartigen Verfahren hergestellt werden, finden beispielsweise in der Nahrungsmittelindustrie Verwendung. So dienen α -Amylasen beispielsweise dazu, um das Altbackenwerden von Brot zu verhindern oder um Fruchtsäfte zu klären. Proteasen werden zum Aufschluß von Proteinen verwendet. All diese Enzyme sind für den Einsatz in Wasch- und Reinigungsmitteln beschrieben, wobei insbesondere die von grampositiven Bakterien bereits natürlicherweise hergestellten Subtilisin-Proteasen einen prominenten Platz einnehmen. Insbesondere in der Textil- und Lederindustrie dienen sie der Aufarbeitung der natürlichen Rohstoffe. Ferner können all diese Enzyme wiederum im Sinne der Biotransformation als Katalysatoren für chemische Reaktionen eingesetzt werden.

Wie einleitend in der Aufgabe formuliert war es erwünscht, ein solches Sicherheitssystem zu finden, das nicht so speziell ist, daß es nicht auch in anderen molekularbiologischen

Ansätzen Verwendung finden könnte. Derartige Ansätze können zu einem weiteren eigenständigen Erfindungsgegenstand zusammengefaßt werden.

Das bedeutet allgemein formuliert die Verwendung des oben beschriebenen Faktors RecA in einem molekularbiologischen Reaktionsansatz. Hierfür werden dessen natürlicherweise vorhandenen Aktivitäten ausgenutzt.

Dementsprechend bevorzugt ist die Verwendung zum Stabilisieren einzelsträngiger DNA, insbesondere bei einer DNA-Polymerisation, bei in vitro erfolgenden Rekombinationsvorgängen, oder zum Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt.

Denn RecA ist ein DNA-einzelstrangbindendes Protein, welches wie erläutert auch eine gewisse Affinität zu doppelsträngiger DNA aufweist. Bei dem natürlichen Prozeß des Crossing over im Zuge der homologen Rekombination kommt diese Funktion zum Tragen. So kann RecA beispielsweise einer PCR oder einer Präparation von Phagen-DNA zugegeben werden, um die Einzelstränge zu stabilisieren. Wenn in vitro Rekombinationsvorgänge nachvollzogen werden, etwa beim Einführen von Mutationen (so auch bei den oben ausgeführten erfindungsgemäßen Mutationen), kann dies von RecA unterstützt werden. Schließlich ist unter dem Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt eine Gyrase- oder Gyrase-unterstützende Funktion gemeint. Dies kann für die Beeinflussung der DNA-Topologie, etwa bei der Arbeit mit Plasmid-DNA ausgenutzt werden.

Einen eigenen Erfindungsgegenstand stellen Vektoren dar, die eine zuvor beschriebene Nukleinsäure enthalten. Denn auch in dieser Form wird die vorliegende Erfindung verwirklicht. So kann diese DNA in Form von Klonierungsvektoren molekularbiologisch bearbeitet oder gelagert werden.

Vorzugsweise handelt es sich bei einem solchen Vektor um einen Expressionsvektor. Denn dieser kann dazu ausgenutzt werden, um ein erfindungsgemäßes RecA herzustellen und den genannten Anwendungen des Faktors zuzuführen.

Einen eigenen Erfindungsgegenstand stellen dementsprechend auch Verfahren zur Herstellung eines zuvor beschriebenen Faktors RecA dar.

Vorzugsweise handelt es sich dabei um Verfahren, die unter Einsatz einer der oben beschriebenen Nukleinsäuren, insbesondere solchen mit zunehmenden Hologiewerten zur SEQ ID NO. 1 erfolgen, vorzugsweise eines entsprechenden Expressionsvektors und weiter bevorzugt durch Fermentation eines dieser Nukleinsäure beziehungsweise diesen Expressionsvektor enthaltenden Wirts.

So wird die vorliegende Erfindung dadurch verwirklicht, daß eine Zelle ein solches Gen in Form einer chromosomalen Kopie erhält und translatiert. Leichter steuerbar erscheint demgegenüber die Bereitstellung dieses Gens in Form eines Plasmids, das gegebenenfalls in mehreren Kopien für die Bildung dieses Faktors sorgt.

Einen eigenen Erfindungsgegenstand stellt die Verwendung der für einen Faktor RecA codierenden Nukleinsäure zur Expression dieses Faktors dar. Denn entsprechend dem oben Gesagten wird dadurch die vorliegende Erfindung zumindest in einem Aspekt verwirklicht.

Vorzugsweise dient das dazu, um diesen Faktor selbst herzustellen, insbesondere in einem oben beschriebenen Verfahren. Alternativ hierzu kann die intrazelluläre Expression auch dazu dienen, um molekularbiologische Aktivitäten der betreffenden Zellen zu modulieren, insbesondere bei in vivo erfolgenden Rekombinationsvorgängen.

Hierbei ist beispielsweise an die Inaktivierung durch einen Antisense- oder RNA-Interferenz-Ansatz gedacht, nach welchem die für RecA codierende mRNA gezielt ausgeschaltet oder nur in einem Teil translatierbar macht. Hierdurch kann die Expression dieses Faktors ganz gezielt moduliert werden. Das gilt sowohl für biotechnologische Produktionsstämme als auch für Laboransätze zum Studium molekularbiologischer Aspekte.

Hierzu gehört schließlich auch eine derartige Verwendung der für einen Faktor RecA codierenden, oben beschriebenen Nukleinsäure zur Inaktivierung dieses Faktors oder des Gens recA in einem in vitro-Ansatz, insbesondere über Wechselwirkung mit einer zugehörigen Nukleinsäure. Das kann besonders für in-vitro-Transkriptions- oder- Translationsansätze vorteilhaft sein, um Rekombinationsvorgänge zu unterbinden.

Beschreibung der Figuren

Figur 1: Aminosäuresequenz-Alignment von SEQ ID NO. 2 mit denen der im Stand der Technik beschriebenen nächstähnlichen Rec-Faktoren.

Dabei bedeuten:

- 1: Faktor RecA aus *B. licheniformis* DSM 13 (SEQ ID NO. 2)
- 2: Faktor RecA aus *B. amyloliquefaciens* (AJ515542 in NCBI)
- 3: Faktor RecA aus *B. subtilis* (Z99112 in NCBI; Region 161035 bis 162078)
- 4: Faktor RecE aus *B. subtilis* (X52132 in NCBI)

Figur 2: Nukleinsäuresequenz-Alignment von SEQ ID NO. 1 mit denen der im Stand der Technik beschriebenen nächstähnlichen rec-Gene.

Dabei bedeuten:

- 1: Gen recA aus *B. licheniformis* DSM 13 (SEQ ID NO. 1)
- 2: Gen recA aus *B. amyloliquefaciens* (AJ515542 in NCBI)
- 3: Gen recA aus *B. subtilis* (Z99112 in NCBI; Region 161035 bis 162078)
- 4: Gen recE aus *B. subtilis* (X52132 in NCBI)

Patentansprüche

1. Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz mindestens zu 96% identisch ist.
2. Faktor nach Anspruch 1 mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz zunehmend bevorzugt mindestens zu 96,5%, 97%, 97,5%, 98%, 98,5%, 99%, 99,5% und ganz besonders bevorzugt zu 100% identisch ist.
3. Faktor RecA, codiert von einer Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.
4. Faktor nach Anspruch 3, codiert von einer Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.
5. Für einen Faktor RecA codierende Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.
6. Nukleinsäure nach Anspruch 6, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.
7. Nukleinsäure nach Anspruch 5 oder 6, codierend für einen Faktor RecA nach einem der Ansprüche 1 bis 4.
8. Verwendung einer für einen Faktor RecA codierenden Nukleinsäure zur funktionellen Inaktivierung des Gens *recA* in einem grampositiven Bakterium, welches nicht *Bacillus megaterium* ist.

9. Verwendung nach Anspruch 8, wobei eine für ein nichtaktives Protein codierende Nukleinsäure mit einer Punktmutation eingesetzt wird.
10. Verwendung nach Anspruch 8, wobei eine Nukleinsäure mit einer Deletions- oder Insertionsmutation eingesetzt wird, vorzugsweise umfassend die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassenden Randsequenzen des für das Protein codierenden Bereichs.
11. Verwendung nach Anspruch 8, wobei Nukleinsäuren mit insgesamt zwei Nukleinsäureabschnitten eingesetzt werden, die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassen und damit den für das Protein codierenden Bereich zumindest teilweise, vorzugsweise vollständig flankieren.
12. Verwendung nach einem der Ansprüche 8 bis 11, wobei es sich um eine Nukleinsäure nach einem der Ansprüche 5 bis 7 handelt, beziehungsweise um die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren.
13. Verwendung nach einem der Ansprüche 8 bis 12, wobei das grampositive Bakterium, vorzugsweise eines der Gattungen Clostridium oder Bacillus, natürlicherweise zur Sporulation befähigt ist und gleichzeitig mit recA ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.
14. Verwendung nach Anspruch 13, wobei es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von B. subtilis um eines der Gene spoIV_A, spoIV_B, spoIV_CA, spoIV_CB, spoIV_FA, spoIV_FB oder yqfD beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von B. subtilis um das Gen yqfD, im Fall von Bacillus licheniformis um das Gen spoIV und in jedem anderen Fall um ein hierzu homologes Gen.
15. Verwendung nach Anspruch 13 oder 14, wobei genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.
16. Verwendung nach Anspruch 14 oder 15, wobei die funktionelle Inaktivierung der Gene spoIV_A, spoIV_B, spoIV_CA, spoIV_CB, spoIV_FA, spoIV_FB, yqfD oder spoIV

beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

17. Grampositives Bakterium, welches nicht *Bacillus megaterium* ist, bei dem das Gen *recA* funktionell inaktiviert ist.
18. Grampositives Bakterium nach Anspruch 17, wobei die funktionelle Inaktivierung über Punktmutagenese, teilweise Deletion oder Insertion oder vollständige Deletion des für das vollständige Protein codierenden Bereichs erfolgt ist.
19. Grampositives Bakterium nach Anspruch 17 oder 18, wobei die funktionelle Inaktivierung über eine Nukleinsäure nach einem der Ansprüche 5 bis 7 beziehungsweise über die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren erfolgt ist.
20. Grampositives Bakterium nach einem der Ansprüche 17 bis 19, vorzugsweise eines der Gattungen *Clostridium* oder *Bacillus*, das natürlicherweise zur Sporulation befähigt ist und bei dem gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.
21. Grampositives Bakterium nach Anspruch 20, wobei es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spolVA*, *spolVB*, *spolVCA*, *spolVCB*, *spolVFA*, *spolVFB* oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spolV* und in jedem anderen Fall um ein hierzu homologes Gen.
22. Grampositives Bakterium nach Anspruch 20 oder 21, wobei genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.

23. Grampositives Bakterium nach Anspruch 21 oder 22, wobei die funktionelle Inaktivierung der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB*, *yqfD* oder *spoIV* beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt ist, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

24. Grampositives Bakterium nach einem der Ansprüche 17 bis 23, wobei es sich um eines der Gattungen Clostridium oder Bacillus handelt, insbesondere um eines der Spezies *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. stearothermophilus*, *B. globigii*, *B. clausii* oder *B. latus*, und ganz besonders um einen Stamm von *B. licheniformis*.

25. Verfahren zur Fermentation eines grampositiven Bakteriums nach einem der Ansprüche 17 bis 24.

26. Verfahren nach Anspruch 25 zur Herstellung eines Wertstoffs, insbesondere einer niedermolekularen Verbindung oder eines Proteins.

27. Verfahren nach Anspruch 26, wobei es sich bei der niedermolekularen Verbindung um einen Naturstoff, einen Nahrungsmittelergänzungsstoff oder um eine pharmazeutisch relevante Verbindung handelt.

28. Verfahren nach Anspruch 26, wobei es sich bei dem Protein um ein Enzym handelt, insbesondere eines aus der Gruppe der α -Amylasen, Proteasen, Cellulasen, Lipasen, Oxidoreduktasen, Peroxidasen, Laccasen, Oxidasen und Hemicellulasen.

29. Verwendung des Faktors RecA nach einem der Ansprüche 1 bis 4 in einem molekularbiologischen Reaktionsansatz.

30. Verwendung nach Anspruch 29 zum Stabilisieren einzelsträngiger DNA, insbesondere bei einer DNA-Polymerisation, bei in vitro erfolgenden Rekombinationsvorgängen, oder zum Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt.

31. Vektor, enthaltend eine Nukleinsäure nach einem der Ansprüche 5 bis 7.
32. Vektor nach Anspruch 31, wobei es sich um einen Expressionsvektor handelt.
33. Verfahren zur Herstellung eines Faktors RecA nach einem der Ansprüche 1 bis 4.
34. Verfahren nach Anspruch 33, unter Einsatz einer Nukleinsäure nach einem der Ansprüche 5 bis 7, vorzugsweise eines Expressionsvektors nach Anspruch 32, weiter bevorzugt durch Fermentation eines diese Nukleinsäure beziehungsweise diesen Expressionsvektor enthaltenden Wirts.
35. Verwendung der für einen Faktor RecA codierenden Nukleinsäure zur Expression dieses Faktors.
36. Verwendung nach Anspruch 35, um diesen Faktor selbst herzustellen, insbesondere in einem Verfahren nach Anspruch 34, oder um molekularbiologische Aktivitäten der betreffenden Zellen zu modulieren, insbesondere bei in vivo erfolgenden Rekombinationsvorgängen.
37. Verwendung der für einen Faktor RecA codierenden Nukleinsäure nach einem der Ansprüche 5 bis 7 zur Inaktivierung dieses Faktors oder des Gens recA in einem in vitro-Ansatz, insbesondere über Wechselwirkung mit einer zugehörigen Nukleinsäure.

Figur 1

1 50
 1 MSDRQAALDM ALKQIEKQFG KGSIMKLGEQ UEURISUVPS GSLALDAALG
 2 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG
 3 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG
 4 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG

51 100
 1 VGGYPRGRII EVYGPESSGK UUVALHAI AE VQQQGGQAAF IDADUALDPV
 2 IGGYPRGRII EVYGPESSGK TTVALHAI AE VQEKGQAAF IDAEHALDPV
 3 IGGYPRGRII EVYGPESSGK TTVALHAI AE VQQQR.TSAF IDAEHALDPV
 4 IGGYPRGRII EVYGPESSGK TTVALHAI AE VQQQR.TSAF IDAEHALDPV

101 150
 1 YAQKLGVNID ELLSQPDUG EQALEIAEAL VRSGAVDIVV IDSVAALVPK
 2 YAQKLGVNIE ELLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK
 3 YAQKLGVNIE ELLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK
 4 YAQKLGVNIE ELLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK

151 200
 1 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKUIAIFI NQIREKVGVM
 2 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM
 3 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM
 4 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM

201 250
 1 FGNPEUUPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKU KIKVVKNKVA
 2 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT RIKVVKNKVA
 3 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT KIKVVKNKVA
 4 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT KIKVVKNKVA

251 300
 1 PPFRTAEVDI MYGEGISKEG EIIDLGUELD IVQKSGAWYS YQEERLGQGR
 2 PPFRTAEVDI MYGEGISKEG EIIDLGTELD IVQKSGSWYS YEEERLGQGR
 3 PPFRTAEVDI MYGEGISKEG EIIDLGTELD IVQKSGSWYS YEEERLGQGR
 4 PPFRTAEVDI MYGEGISKEG EIIDLGTELD IVQKSGSWYS YEEERLGQGR

301 350
 1 ENAKQFLKEN KDILLMIQEQQ IREHYGLDUG GAAPAQEDEA QAQEELEF.S
 2 ENAKQFLKEN KDIIMLMIQEQQ IREHYGLDNN G...VTEKAE EVQEELEFEE
 3 ENAKQFLKEN KDIIMLMIQEQQ IREHYGLDNN G..VVQQQAE ETQEELEFEE
 4 ENAKQFLKEN KDIIMLMIQEQQ IREHYGLDNN G..VVQQQAE ETQEELEFEE

H 06291

Figur 2 / Teil I

1 50
1 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCGCTTAAAC AAATAGAAAA
2 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAC AAATAGAAAA
3 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAC AAATAGAAAA
4 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAC AAATAGAAAA

51 100
1 GCAGTTGGT AAAGGTTCGA TTATGAAACT CGGCGAACAA ACTGAAACGA
2 ACAATTCCGC AAAGGTTCCA TCATGAAGCT CGGAGAAAAA ACGGATACAA
3 ACAGTTCCGC AAAGGTTCCA TTATGAAACT GGGAGAAAAG ACAGATACAA
4 ACAGTTCCGC AAAGGTTCCA TTATGAAACT GGGAGAAAAG ACAGATACAA

101 150
1 GAATTTCAAC AGTTCCGAGC GGTTCTTAG CGCTCGATGC GGCTCTTGG
2 GAATTTCAAC GGTGCCGAGC GGTTCCCTTG CACTTGATAC CGCTCTCGGA
3 GAATTTCTAC TGTACCAAGC GGCTCCCTCG CTCTTGATAC AGCACTGGGA
4 GAATTTCTAC TGTACCAAGC GGCTCCCTCG CTCTTGATAC AGCACTGGGA

151 200
1 GTGGGCGGAT ACCCGCGCGG CCGGATTATT GAAGTATAACG GGCCTGAAAG
2 ATAGGC GGAT ACCCGCGCGG ACGGATTATT GAAGTATAACG GACCTGAAAG
3 ATTGGCGGAT ATCCTCGCGG ACGGATTATT GAAGTATAACG GTCCCTGAAAG
4 ATTGGCGGAT ATCCTCGCGG ACGGATTATT GAAGTATAACG GTCCCTGAAAG

201 250
1 CTCCGGTAAA ACGACGGTGG CGCTTCATGC GATTGCCGAA GTTCAGCAGC
2 CTCAGGTAAA ACGACTGTAG CGCTTCACGC AATCGCTGAG GTTCAGGAAA
3 CTCAGGTAAA ACAACTGTGG CGCTTCATGC GATTGCTGAA GTTCAGCAGC
4 CTCAGGTAAA ACAACTGTGG CGCTTCATGC GATTGCTGAA GTTCAGCAGC

251 300
1 AGGGCGGACA AGCGGCGTTC ATCGACGCCG ACACCGCGCT TGATCCCGTC
2 AAGGC GGACA GGCAGCATT ATTGATGCCG AGCATGCTCT TGATCCTGTG
3 A.. GCGGACA AGC.GCGTT ATCGATGCCG AGCATGCGTT AGATCCGGTA
4 A.. GCGGACA AGC.GCGTT ATCGATGCCG AGCATGCGTT AGATCCGGTA

301 350
1 TATGCACAAA AGCTGGCGT CAACATTGAT GAGCTTTGC TGTACAGCC
2 TACGCGCAAA AGCTCGGTGT CAATATCGAA GAGCTGCTGC TTTCTCAGCC
3 TACGCGCAAA AGCTCGGTGT TAACATCGAA GAGCTTTAC TGTCTCAGCC
4 TACGCGCAAA AGCTCGGTGT TAACATCGAA GAGCTTTAC TGTCTCAGCC

Figur 2 / Teil II

351 400

1 TGATACGGGC GAGCAGGCAG TCGAAATCGC TGAAGCCCTT GTCAGAAGCG
 2 GGATACGGGA GAGCAGGCAG TGGAGATTGC TGAAGCGCTG GTGCGAAGCG
 3 TGACACAGGC GAGCAGGCAG TTGAAATTGC GGAAGCATTG GTTCGAAGCG
 4 TGACACAGGC GAGCAGGCAG TTGAAATTGC GGAAGCATTG GTTCGAAGCG

401 450

1 GAGCGGTGGA TATCGTTGTC ATCGACTCTG TAGCAGCGCT TGTGCCGAAA
 2 GAGCTGTCGA TATCGTAGTC GTTGACTCTG TTGCGGCGCT TGTTCAAAA
 3 GGGCAGTTGA CATTGTCGTT GTCGACTCTG TAGCCGCTCT CGTTCCGAAA
 4 GGGCAGTTGA CATTGTCGTT GTCGACTCTG TAGCCGCTCT CGTTCCGAAA

451 500

1 GCTGAAATCG AAGGAGATAT GGGGGATTCC CACGTCGGTT TGCAGGCCAG
 2 GCTGAAATTG AAGGTGACAT GGGTGATTCA CACGTCGGTT TACAGGCCAG
 3 GCGGAAATTG AAGGCGACAT GGGAGATTAG CATGTCGGTT TACAAGCACG
 4 GCGGAAATTG AAGGCGACAT GGGAGATTAG CATGTCGGTT TACAAGCACG

501 550

1 ACTGATGTCT CAGGCCTTC GCAAGCTTC CGGAGCGATC AATAAATCGA
 2 TCTCATGTCT CAGGCCTCC GTAAGCTTC CGGCGCCATC AATAAATCTA
 3 CTTAATGTCT CAAGCGCTTC GTAAGCTTC AGGGGCCATT AACAAATCGA
 4 CTTAATGTCT CAAGCGCTTC GTAAGCTTC AGGGGCCATT AACAAATCGA

551 600

1 AGACCATCGC GATCTTATC AACCAAGATTC GTGAAAAAGT CGGTGTCATG
 2 AAACAATCGC AATCTTATT AACCAAGATTC GTGAAAAAGT CGGCCTTATG
 3 AGACAATCGC GATTTTCATT AACCAAATTC GTGAAAAAGT CGGTGTTATG
 4 AGACAATCGC GATTTTCATT AACCAAATTC GTGAAAAAGT CGGTGTTATG

601 650

1 TTTGGAAATC CTGAGACGAC GCCAGCGGA AGAGCGCTGA AATTCTACTC
 2 TTCCGAAATC CGGAGACGAC ACCGGCGGC CGCGCGCTGA AATTCTATTC
 3 TTCCGGAAACC CGGAAACAAAC ACCTGGCGGC CGTGCCTTGA AATTCTATTC
 4 TTCCGGAAACC CGGAAACAAAC ACCTGGCGGC CGTGCCTTGA AATTCTATTC

651 700

1 TTCTGTCCGC CTTGAAGTGC GCCCGCGAGA GCAGCTGAAA CAAGGCAACG
 2 TTCCGTGCGT CTTGAAGTGC GCCGTGCCGA GCAATTAAAG CAGGGCAACG
 3 TTCCGTGCGT CTTGAAGTGC GCCGTGCTGA ACAGCTGAAA CAAGGCAACG
 4 TTCCGTGCGT CTTGAAGTGC GCCGTGCTGA ACAGCTGAAA CAAGGCAACG

Figur 2 / Teil III

	701	750			
1	ACGT CATGGG	GAACAAGACG	AAAATCAAAG	TCGTGAAAAA	CAAAGTGGCA
2	ACGTTATGGG	GAATAAAACG	AGAATTAAG	TCGTAAAAAA	CAAAGTCGCT
3	ACGTAATGGG	GAACAAAACG	AAAATCAAAG	TCGTGAAAAA	CAAGGTGGCT
4	ACGTAATGGG	GAACAAAACG	AAAATCAAAG	TCGTGAAAAA	CAAGGTGGCT
	751	800			
1	CCTCCATTCC	GGACAGCCGA	AGTGGACATT	ATGTACGGGG	AAGGAATTTC
2	CCTCCGTTCC	GTACGGCTGA	AGTGGACATT	ATGTACGGTG	AAGGAATCTC
3	CCGCCGTTCC	GTACAGCCGA	GGTTGACATT	ATGTACGGAG	AAGGCATTT
4	CCGCCGTTCC	GTACAGCCGA	GGTTGACATT	ATGTACGGAG	AAGGCATTT
	801	850			
1	AAAAGAAGGG	GAAATCATCG	ACCTCGGAAC	AGAGCTTGAC	ATCGTTCAAA
2	CAAAGAAGGG	GAAATCATCG	ACCTTGGAAC	TGAACTTGAT	ATCGTGCAGA
3	AAAAGAAGGC	GAAATCATTG	ATCTAGGAAC	TGAACTTGAT	ATCGTGCAAA
4	AAAAGAAGGC	GAAATCATTG	ATCTAGGAAC	TGAACTTGAT	ATCGTGCAAA
	851	900			
1	AGAGCGGTGC	ATGGTACTCT	TATCAGGAGG	AACGCCTTGG	ACAAGGCCGT
2	AAAGCGGCTC	GTGGTATTCT	TATGAAGAAG	AACGCCTCGG	ACAGGGCCGT
3	AAAGCGGTTTC	ATGGTACTCT	TATGAAGAAG	AGCGTCTTGG	CCAAGGCCGT
4	AAAGCGGTTTC	ATGGTACTCT	TATGAAGAAG	AGCGTCTTGG	CCAAGGCCGT
	901	950			
1	GAAAACGCCA	AACAGTTCCCT	GAAAGAAAAC	AAGGATATCC	TTTTGATGAT
2	GAAAACGCCA	AGCAGTTCTT	AAAAGAAAAT	AAAGACATCA	TGCTGATGAT
3	GAAAATGCAA	AACAATTCCCT	GAAAGAAAAT	AAAGATATCA	TGCTGATGAT
4	GAAAATGCAA	AACAATTCCCT	GAAAGAAAAT	AAAGATATCA	TGCTGATGAT
	951	1000			
1	TCAAGAGCAG	ATCCGGGAGC	ACTACGGTTT	GGATACTGGA	GGCGCTGCTC
2	TCAAGAACAA	ATCCGTGAAC	ATTACGGTTT	GGACAATAAC	GGTGTAC..
3	CCAGGAGCAA	ATTCGCGAAC	ATTACGGCTT	GGATAATAAC	GGAGTAGTGC
4	CCAGGAGCAA	ATTCGCGAAC	ATTACGGCTT	GGATAATAAC	GGAGTAGTGC
	1001	1050			
1	CTGCACAGGA	AGACGAGGCC	CAAGCTCAGG	AAGAACTCGA	GTTTTAATCA
2GGA	AAAAGCGGAA	GAAGTTCAAG	AAGAGCTTGA	ATTCGAAGAA
3AGCA	GCAAGCTGAA	GAGACACAAG	AAGAACTCGA	ATTGAAAGAA
4AGCA	GCAAGCTGAA	GAGACACAAG	AAGAACTCGA	ATTGAAAGAA
	1051				
1	TGA				
2	TAA				
3	...				
4	TAA				

Zusammenfassung

Gegenstand der vorliegenden Anmeldung ist der Faktor RecA aus *Bacillus licheniformis* DSM 13 (SEQ ID NO. 2), zusammen mit dem zugehörigen Gen recA (SEQ ID NO. 1), einschließlich verwandten Proteinen und Genen hierzu. Das Gen recA wird erfindungsgemäß unter anderem zur Konstruktion von grampositiven bakteriellen Sicherheitsstämmen für die biotechnologische Produktion eingesetzt, indem es in den betreffenden Stämmen inaktiviert wird. Diese weisen, sofern sie natürlicherweise zur Sporulation befähigt sind, in einer speziellen Ausführungsform zusätzliche funktionelle Deletionen in Phase-IV-Sporulationsgenen auf, vorzugsweise im Gen *spolV* (bei *Bacillus licheniformis*), im Gen *yqfD* (bei *B. subtilis*) beziehungsweise in dem hierzu jeweils homologen Gen. Des weiteren steht mit diesem RecA ein Protein zur Verfügung, das in molekularbiologischen Ansätzen oder zur Modulation der molekularbiologischen Aktivitäten von Zellen gebraucht werden kann, insbesondere im Zusammenhang mit DNA-Polymerisations- oder Rekombinationsvorgängen.

H 06291

SEQUENZPROTOKOLL

<110> Henkel Kommanditgesellschaft auf Aktien
<120> Der Faktor RecA aus *Bacillus licheniformis* und
recA-inaktivierte Sicherheitsstämme für die
biotechnologische Produktion
<130> H 06291 DE
<140>
<141>
<160> 18
<170> PatentIn Ver. 2.1
<210> 1
<211> 1047
<212> DNA
<213> *Bacillus licheniformis* DSM 13
<220>
<221> CDS
<222> (1)..(1047)
<220>
<221> gene
<222> (1)..(1047)
<223> recA
<400> 1
atg agt gat cgt cag gca gcc tta gat atg gcg ctt aaa caa ata gaa 48
Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln Ile Glu
1 5 10 15
aag cag ttt ggt aaa ggt tcg att atg aaa ctc ggc gaa caa act gaa 96
Lys Gln Phe Gly Lys Gly Ser Ile Met Lys Leu Gly Glu Gln Thr Glu
20 25 30
acg aga att tca aca gtt ccg agc ggt tct tta gcg ctc gat gcg gct 144
Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp Ala Ala
35 40 45
ctt gga gtg ggc gga tac ccg cgc ggc cg att att gaa gta tac ggg 192
Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val Tyr Gly
50 55 60
cct gaa agc tcc ggt aaa acg acg gtg gcg ctt cat gcg att gcc gaa 240
Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile Ala Glu
65 70 75 80
gtt cag cag cag ggc gga caa gcg gcg ttc atc gac gcc gac acc gcg 288
Val Gln Gln Gln Gly Gly Gln Ala Ala Phe Ile Asp Ala Asp Thr Ala
85 90 95
ctt gat ccc gtc tat gca caa aag ctg ggc gtc aac att gat gag ctt 336
Leu Asp Pro Val Tyr Ala Gln Lys Leu Gly Val Asn Ile Asp Glu Leu
100 105 110

ttg ctg tca cag cct gat acg ggc gag cag gcg ctc gaa atc gct gaa	384
Leu Leu Ser Gln Pro Asp Thr Gly Glu Gln Ala Leu Glu Ile Ala Glu	
115 120 125	
gcc ctt gtc aga agc gga gcg gtg gat atc gtt gtc atc gac tct gta	432
Ala Leu Val Arg Ser Gly Ala Val Asp Ile Val Val Ile Asp Ser Val	
130 135 140	
gca gcg ctt gtg ccg aaa gct gaa atc gaa gga gat atg ggg gat tcc	480
Ala Ala Leu Val Pro Lys Ala Glu Ile Glu Gly Asp Met Gly Asp Ser	
145 150 155 160	
cac gtc ggt ttg cag gcc aga ctg atg tct cag gcg ctt cgc aag ctt	528
His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg Lys Leu	
165 170 175	
tcc gga gcg atc aat aaa tcg aag acc atc gcg atc ttt atc aac cag	576
Ser Gly Ala Ile Asn Lys Ser Lys Thr Ile Ala Ile Phe Ile Asn Gln	
180 185 190	
att cgt gaa aaa gtc ggt gtc atg ttt gga aat cct gag acg acg cca	624
Ile Arg Glu Lys Val Gly Val Met Phe Gly Asn Pro Glu Thr Thr Pro	
195 200 205	
ggc gga aga gcg ctg aaa ttc tac tct tct gtc cgc ctt gaa gtg cgc	672
Gly Gly Arg Ala Leu Lys Phe Tyr Ser Ser Val Arg Leu Glu Val Arg	
210 215 220	
cgc gca gag cag ctg aaa caa ggc aac gac gtc atg ggg aac aag acg	720
Arg Ala Glu Gln Leu Lys Gln Gly Asn Asp Val Met Gly Asn Lys Thr	
225 230 235 240	
aaa atc aaa gtc gtg aaa aac aaa gtg gca cct cca ttc cgg aca gcc	768
Lys Ile Lys Val Val Lys Asn Val Ala Pro Pro Phe Arg Thr Ala	
245 250 255	
gaa gtg gac att atg tac ggg gaa gga att tca aaa gaa ggg gaa atc	816
Glu Val Asp Ile Met Tyr Gly Glu Gly Ile Ser Lys Glu Gly Glu Ile	
260 265 270	
atc gac ctc gga aca gag ctt gac atc gtt caa aag acg ggt gca tgg	864
Ile Asp Leu Gly Thr Glu Leu Asp Ile Val Gln Lys Ser Gly Ala Trp	
275 280 285	
tac tct tat cag gag gaa cgc ctt gga caa ggc cgt gaa aac gcc aaa	912
Tyr Ser Tyr Gln Glu Glu Arg Leu Gly Gln Gly Arg Glu Asn Ala Lys	
290 295 300	
cag ttc ctg aaa gaa aac aag gat atc ctt ttg atg att caa gag cag	960
Gln Phe Leu Lys Glu Asn Lys Asp Ile Leu Leu Met Ile Gln Glu Gln	
305 310 315 320	
atc cgg gag cac tac ggt ttg gat act gga ggc gct gct cct gca cag	1008
Ile Arg Glu His Tyr Gly Leu Asp Thr Gly Gly Ala Ala Pro Ala Gln	
325 330 335	
gaa gac gag gcc caa gct cag gaa gaa ctc gag ttt taa	1047
Glu Asp Glu Ala Gln Ala Glu Glu Leu Glu Phe	

<210> 2
 <211> 348
 <212> PRT
 <213> *Bacillus licheniformis* DSM 13

<400> 2
 Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln Ile Glu
 1 5 10 15
 Lys Gln Phe Gly Lys Gly Ser Ile Met Lys Leu Gly Glu Gln Thr Glu
 20 25 30
 Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp Ala Ala
 35 40 45
 Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val Tyr Gly
 50 55 60
 Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile Ala Glu
 65 70 75 80
 Val Gln Gln Gln Gly Gly Gln Ala Ala Phe Ile Asp Ala Asp Thr Ala
 85 90 95
 Leu Asp Pro Val Tyr Ala Gln Lys Leu Gly Val Asn Ile Asp Glu Leu
 100 105 110
 Leu Leu Ser Gln Pro Asp Thr Gly Glu Gln Ala Leu Glu Ile Ala Glu
 115 120 125
 Ala Leu Val Arg Ser Gly Ala Val Asp Ile Val Val Ile Asp Ser Val
 130 135 140
 Ala Ala Leu Val Pro Lys Ala Glu Ile Glu Gly Asp Met Gly Asp Ser
 145 150 155 160
 His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg Lys Leu
 165 170 175
 Ser Gly Ala Ile Asn Lys Ser Lys Thr Ile Ala Ile Phe Ile Asn Gln
 180 185 190
 Ile Arg Glu Lys Val Gly Val Met Phe Gly Asn Pro Glu Thr Thr Pro
 195 200 205
 Gly Gly Arg Ala Leu Lys Phe Tyr Ser Ser Val Arg Leu Glu Val Arg
 210 215 220
 Arg Ala Glu Gln Leu Lys Gln Gly Asn Asp Val Met Gly Asn Lys Thr
 225 230 235 240
 Lys Ile Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Arg Thr Ala
 245 250 255
 Glu Val Asp Ile Met Tyr Gly Glu Gly Ile Ser Lys Glu Gly Glu Ile
 260 265 270
 Ile Asp Leu Gly Thr Glu Leu Asp Ile Val Gln Lys Ser Gly Ala Trp
 275 280 285
 Tyr Ser Tyr Gln Glu Glu Arg Leu Gly Gln Gly Arg Glu Asn Ala Lys
 290 295 300
 Gln Phe Leu Lys Glu Asn Lys Asp Ile Leu Leu Met Ile Gln Glu Gln
 305 310 315 320
 Ile Arg Glu His Tyr Gly Leu Asp Thr Gly Gly Ala Ala Pro Ala Gln
 325 330 335
 Glu Asp Glu Ala Gln Ala Gln Glu Glu Leu Glu Phe
 340 345

<210> 3
 <211> 1792
 <212> DNA

H 06291

<213> *Bacillus licheniformis*

<220>

<221> CDS

<222> (140)..(1336)

<220>

<221> gene

<222> (1)..(1792)

<223> spoIV

<220>

<221> misc_feature

<222> (140)..(142)

<223> First codon translated as Met.

<400> 3

ggctgatgct caaacagggg cagtgcata ttcaaggcaa agactttgtc atcaaaaacga 60

ttttgcctga ggaaattctg cttgaaggca cgattgagct tgtccgctat atcgattcat 120

aagtcggggg gaaagaagc gtg aag aat aaa tgg ctt tct ttt tca gga 172
Val Lys Asn Lys Trp Leu Ser Phe Phe Ser Gly
1 5 10

aag atc cag ctt aag ata acg gga aaa ggg atc gaa cgg tta tta aat 220
Lys Ile Gln Leu Lys Ile Thr Gly Lys Gly Ile Glu Arg Leu Leu Asn
15 20 25

gaa tgc acc agg cgc aac atc ccg atg ttt aat gta aag aaa aag aaa 268
Glu Cys Thr Arg Arg Asn Ile Pro Met Phe Asn Val Lys Lys Lys Lys
30 35 40

gac gcc gtc ttt ctt tat att ccg ctt tct gat gta cat gcc ttc cgg 316
Asp Ala Val Phe Leu Tyr Ile Pro Leu Ser Asp Val His Ala Phe Arg
45 50 55

aag gtc atc aga ggc ttc gac tgc aag tgc agg ttc atc aaa cga aaa 364
Lys Val Ile Arg Gly Phe Asp Cys Lys Cys Arg Phe Ile Lys Arg Lys
60 65 70 75

ggg ttt cct ttc ctc gtg cag aag tct aaa cgg aat agc ggc ttc act 412
Gly Phe Pro Phe Leu Val Gln Lys Ser Lys Arg Asn Ser Gly Phe Thr
80 85 90

ttt gga gtt gct gca ttt ttt atc atc atg ctc cta ttg tcc aac atg 460
Phe Gly Val Ala Ala Phe Phe Ile Ile Met Leu Leu Ser Asn Met
95 100 105

ctt tgg aaa att gat att aca gga gcc aat ccg gag aca gaa cat caa 508
Leu Trp Lys Ile Asp Ile Thr Gly Ala Asn Pro Glu Thr Glu His Gln
110 115 120

atc aaa cag caa ttg gat caa atc ggc gtc aaa aaa ggc cgc ttt cag 556
Ile Lys Gln Gln Leu Asp Gln Ile Gly Val Lys Lys Gly Arg Phe Gln
125 130 135

ttt tca atg ctg acc ccg gaa aaa att cag cag gcg ctc aca aag cgg 604
Phe Ser Met Leu Thr Pro Glu Lys Ile Gln Gln Ala Leu Thr Lys Arg

H 06291

140	145	150	155	
gtc gaa aac atc act tgg gtg ggt att gag tta aac ggc acc gcc ctt				652
Val Glu Asn Ile Thr Trp Val Gly Ile Glu Leu Asn Gly Thr Ala Leu				
160	165	170		
cac atg aaa gtc gtt gaa aag aat gaa cct gac aaa gaa aaa tat atc				700
His Met Lys Val Val Glu Lys Asn Glu Pro Asp Lys Glu Lys Tyr Ile				
175	180	185		
ggt ccg agg cac atc gtc gcc aaa aaa ggg gcg acc atc tcg aaa aag				748
Gly Pro Arg His Ile Val Ala Lys Lys Gly Ala Thr Ile Ser Lys Lys				
190	195	200		
ttc gtg gaa aaa ggc gag ccg ctc gtc acg gtg aac cag cac gtt gaa				796
Phe Val Glu Lys Gly Glu Pro Leu Val Thr Val Asn Gln His Val Glu				
205	210	215		
aaa ggg caa atg ctc gtt tcc ggg ctg atc gga agc gaa gag gaa aag				844
Lys Gly Gln Met Leu Val Ser Gly Leu Ile Gly Ser Glu Glu Glu Lys				
220	225	230	235	
caa aaa gtc gga gca aaa ggg aaa atc tac ggt gaa acc tgg tac aag				892
Gln Lys Val Gly Ala Lys Gly Lys Ile Tyr Gly Glu Thr Trp Tyr Lys				
240	245	250		
tca aca gta acg gtt cct ctt gag aca tca ttt gac gtt ttt acg ggt				940
Ser Thr Val Thr Val Pro Leu Glu Thr Ser Phe Asp Val Phe Thr Gly				
255	260	265		
aaa gta agg aca agt cac aag cta tcc ctc gga tca ttt tcc gtg ccg				988
Lys Val Arg Thr Ser His Lys Leu Ser Leu Gly Ser Phe Ser Val Pro				
270	275	280		
atc tgg ggc ttt tca ttt aaa aaa gaa gac ttc tcg cgc ccg aag acg				1036
Ile Trp Gly Phe Ser Phe Lys Lys Glu Asp Phe Ser Arg Pro Lys Thr				
285	290	295		
gag acc gaa aac ccc tcg ctg cat ttt atg aat ttt aag ctt cct gtc				1084
Glu Thr Glu Asn Pro Ser Leu His Phe Met Asn Phe Lys Leu Pro Val				
300	305	310	315	
gct tat gaa aag gag cat atg agg gag agc gaa caa atc aaa agg gtg				1132
Ala Tyr Glu Lys Glu His Met Arg Glu Ser Glu Gln Ile Lys Arg Val				
320	325	330		
tac tcg aaa aaa gaa gca gtt ctt gaa gga atc gaa atg gga aaa aga				1180
Tyr Ser Lys Lys Glu Ala Val Leu Glu Gly Ile Glu Met Gly Lys Arg				
335	340	345		
gac atc agg aaa aaa atc ggc agc gac ggg aac att atc agt gaa aaa				1228
Asp Ile Arg Lys Lys Ile Gly Ser Asp Gly Asn Ile Ile Ser Glu Lys				
350	355	360		
gtt ttg cac gaa acg agc gag aat ggc aaa gtt aaa ttg atc atc ctt				1276
Val Leu His Glu Thr Ser Glu Asn Gly Lys Val Lys Leu Ile Ile Leu				
365	370	375		
tac cag gtt att gaa gac att gtt caa aca aca cca att gtt cag gag				1324

H 06291

Tyr Gln Val Ile Glu Asp Ile Val Gln Thr Thr Pro Ile Val Gln Glu
380 385 390 395

act aaa gaa tga cagaacactt acttgcaatt catcagcaac tggaaagtcc 1376
Thr Lys Glu

aatgaggct caaacgctgt ttggAACCA ggattccat ttGAAGTTGA tggAGGAAGA 1436
gctgaacatt tcaattgtca cgcgcggaga aaccgtgtat gtgacaggag atgaagaaac 1496
gtttgaaatc gcggacagcc tgctgcctc tctcctaaat ctgatccgca aaggaatcga 1556
gatatccgaa cgcgatgtct tgtatgcgtat caagatggcg aaaaagcaga agcttgagtt 1616
ttttgaaagc atgtatgaag aggaaattac gaaaaacGCC aaaggaaaac cgatcagagt 1676
caaaaccatc ggtcaaagag aatacatcgc cgccatgaaa aggcacgact taatcttcgg 1736
catcgccccca gcaggaacgg ggaaaaccta tttggctgtc gtaaaggccc ttcatg 1792

<210> 4
<211> 398
<212> PRT
<213> Bacillus licheniformis

<400> 4

Val Lys Asn Lys Trp Leu Ser Phe Phe Ser Gly Lys Ile Gln Leu Lys
1 5 10 15
Ile Thr Gly Lys Gly Ile Glu Arg Leu Leu Asn Glu Cys Thr Arg Arg
20 25 30
Asn Ile Pro Met Phe Asn Val Lys Lys Lys Asp Ala Val Phe Leu
35 40 45
Tyr Ile Pro Leu Ser Asp Val His Ala Phe Arg Lys Val Ile Arg Gly
50 55 60
Phe Asp Cys Lys Cys Arg Phe Ile Lys Arg Lys Gly Phe Pro Phe Leu
65 70 75 80
Val Gln Lys Ser Lys Arg Asn Ser Gly Phe Thr Phe Gly Val Ala Ala
85 90 95
Phe Phe Ile Ile Met Leu Leu Leu Ser Asn Met Leu Trp Lys Ile Asp
100 105 110
Ile Thr Gly Ala Asn Pro Glu Thr Glu His Gln Ile Lys Gln Gln Leu
115 120 125
Asp Gln Ile Gly Val Lys Lys Gly Arg Phe Gln Phe Ser Met Leu Thr
130 135 140
Pro Glu Lys Ile Gln Gln Ala Leu Thr Lys Arg Val Glu Asn Ile Thr
145 150 155 160
Trp Val Gly Ile Glu Leu Asn Gly Thr Ala Leu His Met Lys Val Val
165 170 175
Glu Lys Asn Glu Pro Asp Lys Glu Lys Tyr Ile Gly Pro Arg His Ile
180 185 190
Val Ala Lys Lys Gly Ala Thr Ile Ser Lys Lys Phe Val Glu Lys Gly
195 200 205
Glu Pro Leu Val Thr Val Asn Gln His Val Glu Lys Gly Gln Met Leu
210 215 220
Val Ser Gly Leu Ile Gly Ser Glu Glu Glu Lys Gln Lys Val Gly Ala
225 230 235 240
Lys Gly Lys Ile Tyr Gly Glu Thr Trp Tyr Lys Ser Thr Val Thr Val
245 250 255

Pro Leu Glu Thr Ser Phe Asp Val Phe Thr Gly Lys Val Arg Thr Ser
260 265 270
His Lys Leu Ser Leu Gly Ser Phe Ser Val Pro Ile Trp Gly Phe Ser
275 280 285
Phe Lys Lys Glu Asp Phe Ser Arg Pro Lys Thr Glu Thr Glu Asn Pro
290 295 300
Ser Leu His Phe Met Asn Phe Lys Leu Pro Val Ala Tyr Glu Lys Glu
305 310 315 320
His Met Arg Glu Ser Glu Gln Ile Lys Arg Val Tyr Ser Lys Lys Glu
325 330 335
Ala Val Leu Glu Gly Ile Glu Met Gly Lys Arg Asp Ile Arg Lys Lys
340 345 350
Ile Gly Ser Asp Gly Asn Ile Ile Ser Glu Lys Val Leu His Glu Thr
355 360 365
Ser Glu Asn Gly Lys Val Lys Leu Ile Ile Leu Tyr Gln Val Ile Glu
370 375 380
Asp Ile Val Gln Thr Thr Pro Ile Val Gln Glu Thr Lys Glu
385 390 395

<210> 5
<211> 1594
<212> DNA
<213> *Bacillus subtilis*

<220>
<221> CDS
<222> (201)..(1397)

<220>
<221> gene
<222> (1)..(1594)
<223> *yqfD*

<220>
<221> misc_feature
<222> (201)..(203)
<223> First codon translated as Met.

<400> 5
gacttcataat ctacatagaa aaccacagag gcctttgct tttcagttag aatgaagtgc 60
ggctgatgct gaagcaggcc cagtgcata tatctggtaa aaattttgtc atcaaggcga 120
ttcttccgga agagatactt ttggagggtta cgattgtatgt cgttcgatata gttgagtcata 180
aaagccgagg gggaaatgtt gtg aaa aat aaa tgg ctg tct ttt tcg ggt 233
Val Lys Asn Lys Trp Leu Ser Phe Phe Ser Gly
1 5 10

aag gtc cag ctt gaa ttg acg gga aga ggg att gag cgg ctc ctt aat 281
Lys Val Gln Leu Glu Leu Thr Gly Arg Gly Ile Glu Arg Leu Leu Asn
15 20 25

gaa tgc aca aaa cag ggg att ccg gtc ttt cat gtc aaa aaa aag aaa 329
Glu Cys Thr Lys Gln Gly Ile Pro Val Phe His Val Lys Lys Lys
30 35 40

gaa gcc gta tcg tta tat ata cag ctt cag gat gta cat gcc ttt cg	377
Glu Ala Val Ser Leu Tyr Ile Gln Leu Gln Asp Val His Ala Phe Arg	
45 50 55	
cgg gta aga agt aaa ttt aaa tgt aaa gcc cga ttt atc aat cg aag	425
Arg Val Arg Ser Lys Phe Lys Cys Lys Ala Arg Phe Ile Asn Arg Lys	
60 65 70 75	
gga ttt ccc ttc ctg ttg ctg aaa tca aag, ctg aat ata ggg ttt acg	473
Gly, Phe Pro Phe Leu Leu Lys Ser Lys Leu Asn Ile Gly Phe Thr	
80 85 90	
atc ggt ttt gcg att ttt ttc att ctt ttg ttt ttg ctg tcc aat atg	521
Ile Gly Phe Ala Ile Phe Ile Leu Leu Phe Leu Leu Ser Asn Met	
95 100 105	
gtg tgg aaa att gat gtg aca ggc gct aag cct gaa aca gaa cat caa	569
Val Trp Lys Ile Asp Val Thr Gly Ala Lys Pro Glu Thr Glu His Gln	
110 115 120	
atg agg cag cat ctt aat gaa atc ggc gtc aaa aag ggc cgt ctg cag	617
Met Arg Gln His Leu Asn Glu Ile Gly Val Lys Lys Gly Arg Leu Gln	
125 130 135	
ttt tta atg atg tcg ccc gaa aaa ata cag aaa tca tta acc aat gga	665
Phe Leu Met Met Ser Pro Glu Lys Ile Gln Lys Ser Leu Thr Asn Gly	
140 145 150 155	
ata gac aat atc act tgg gtc gga gtt gat ctg aag ggg acg acc att	713
Ile Asp Asn Ile Trp Val Gly Val Asp Leu Lys Gly Thr Thr Ile	
160 165 170	
cat atg aaa gtt gtg gag aaa aat gag ccc gaa aaa gaa aaa tat gtt	761
His Met Lys Val Val Glu Lys Asn Glu Pro Glu Lys Glu Lys Tyr Val	
175 180 185	
agc ccg cgc aat att gtc gcc aaa aag aaa gca acc att acg aga atg	809
Ser Pro Arg Asn Ile Val Ala Lys Lys Lys Ala Thr Ile Thr Arg Met	
190 195 200	
tct gtg caa aaa gga cag ccc atg gcc gcc ata cac gat cat gtt gaa	857
Ser Val Gln Lys Gly Gln Pro Met Ala Ala Ile His Asp His Val Glu	
205 210 215	
aag gga cag ctg ctt gtt tcg gga ctg atc ggc agc gaa gac cat cag	905
Lys Gly Gln Leu Leu Val Ser Gly Leu Ile Gly Ser Glu Asp His Gln	
220 225 230 235	
cag gaa gtc gcc tca aaa gca gaa att tat gga gaa acc tgg tat aga	953
Gln Glu Val Ala Ser Lys Ala Glu Ile Tyr Gly Glu Thr Trp Tyr Arg	
240 245 250	
tca gaa gtg aca gtc ccg ctt gaa aca tta ttt aac gtc tat acg ggc	1001
Ser Glu Val Thr Val Pro Leu Glu Thr Leu Phe Asn Val Tyr Thr Gly	
255 260 265	
aaa gta agg aca aag cac aag ctt tct ttt ggt tct ttg gca atc ccg	1049
Lys Val Arg Thr Lys His Lys Leu Ser Phe Gly Ser Leu Ala Ile Pro	
270 275 280	

atc tgg ggg atg acg ttt aaa aaa gag gaa ttg aag cat cca aaa aca Ile Trp Gly Met Thr Phe Lys Lys Glu Glu Leu Lys His Pro Lys Thr 285 290 295	1097	
gaa caa gaa aag cat tcg ctt cat ttt ctc gga ttt aag ctc cct gta Glu Gln Glu Lys His Ser Leu His Phe Leu Gly Phe Lys Leu Pro Val 300 305 310 315	1145	
tcc tat gtc aaa gag caa acg aga gaa agt gaa gag gct ttg cga aaa Ser Tyr Val Lys Glu Gln Thr Arg Glu Ser Glu Glu Ala Leu Arg Lys 320 325 330	1193	
tat aca aaa gaa gaa gca gtt caa gaa ggc att aaa ttg ggt aaa cag Tyr Thr Lys Glu Glu Ala Val Gln Glu Gly Ile Lys Leu Gly Lys Gln 335 340 345	1241	
gat gta gag gat aaa ata ggc gaa aac ggc gag gtg aaa agt gaa aaa Asp Val Glu Asp Lys Ile Gly Glu Asn Gly Glu Val Lys Ser Glu Lys 350 355 360	1289	
gtt ttg cac cag act gtt gag aat ggt aaa gta aag ttg att att ctc Val Leu His Gln Thr Val Glu Asn Gly Lys Val Lys Leu Ile Ile Leu 365 370 375	1337	
tac caa gtt ata gaa gat atc gtt caa acc aca cct att gtc agg gag Tyr Gln Val Ile Glu Asp Ile Val Gln Thr Thr Pro Ile Val Arg Glu 380 385 390 395	1385	
act gaa gaa tga cagaacattt acttgcgatg aatcaaaaac tgaaaaaccc Thr Glu Glu	1437	
ggacgaggcg ctttcactct tcggAACCA agattcttt ttgaaattga tggagaaaga 1497		
tctgaattta aatatcatta cgcgcggcga gacgatttat gttcaggcg atgatgaatc 1557		
gtttcagatt gcagacagggc tgctgggatc gctcctc 1594		
<210> 6 <211> 398 <212> PRT <213> <i>Bacillus subtilis</i>		
<400> 6 Val Lys Asn Lys Trp Leu Ser Phe Phe Ser Gly Lys Val Gln Leu Glu 1 5 10 15 Leu Thr Gly Arg Gly Ile Glu Arg Leu Leu Asn Glu Cys Thr Lys Gln 20 25 30 Gly Ile Pro Val Phe His Val Lys Lys Lys Lys Glu Ala Val Ser Leu 35 40 45 Tyr Ile Gln Leu Gln Asp Val His Ala Phe Arg Arg Val Arg Ser Lys 50 55 60 Phe Lys Cys Lys Ala Arg Phe Ile Asn Arg Lys Gly Phe Pro Phe Leu 65 70 75 80 Leu Leu Lys Ser Lys Leu Asn Ile Gly Phe Thr Ile Gly Phe Ala Ile 85 90 95 Phe Phe Ile Leu Leu Phe Leu Leu Ser Asn Met Val Trp Lys Ile Asp 100 105 110		

H 06291

Val Thr Gly Ala Lys Pro Glu Thr Glu His Gln Met Arg Gln His Leu
115 120 125
Asn Glu Ile Gly Val Lys Lys Gly Arg Leu Gln Phe Leu Met Met Ser
130 135 140
Pro Glu Lys Ile Gln Lys Ser Leu Thr Asn Gly Ile Asp Asn Ile Thr
145 150 155 160
Trp Val Gly Val Asp Leu Lys Gly Thr Thr Ile His Met Lys Val Val
165 170 175
Glu Lys Asn Glu Pro Glu Lys Glu Lys Tyr Val Ser Pro Arg Asn Ile
180 185 190
Val Ala Lys Lys Ala Thr Ile Thr Arg Met Ser Val Gln Lys Gly
195 200 205
Gln Pro Met Ala Ala Ile His Asp His Val Glu Lys Gly Gln Leu Leu
210 215 220
Val Ser Gly Leu Ile Gly Ser Glu Asp His Gln Gln Glu Val Ala Ser
225 230 235 240
Lys Ala Glu Ile Tyr Gly Glu Thr Trp Tyr Arg Ser Glu Val Thr Val
245 250 255
Pro Leu Glu Thr Leu Phe Asn Val Tyr Thr Gly Lys Val Arg Thr Lys
260 265 270
His Lys Leu Ser Phe Gly Ser Leu Ala Ile Pro Ile Trp Gly Met Thr
275 280 285
Phe Lys Lys Glu Glu Leu Lys His Pro Lys Thr Glu Gln Glu Lys His
290 295 300
Ser Leu His Phe Leu Gly Phe Lys Leu Pro Val Ser Tyr Val Lys Glu
305 310 315 320
Gln Thr Arg Glu Ser Glu Glu Ala Leu Arg Lys Tyr Thr Lys Glu Glu
325 330 335
Ala Val Gln Glu Gly Ile Lys Leu Gly Lys Gln Asp Val Glu Asp Lys
340 345 350
Ile Gly Glu Asn Gly Glu Val Lys Ser Glu Lys Val Leu His Gln Thr
355 360 365
Val Glu Asn Gly Lys Val Lys Leu Ile Ile Leu Tyr Gln Val Ile Glu
370 375 380
Asp Ile Val Gln Thr Thr Pro Ile Val Arg Glu Thr Glu Glu
385 390 395

<210> 7
<211> 1876
<212> DNA
<213> *Bacillus subtilis*

<220>
<221> CDS
<222> (201)..(1679)

<220>
<221> gene
<222> (1)..(1876)
<223> *spoIVVA*

<220>
<221> misc_feature
<222> (201)..(203)
<223> First codon translated as Met.

<400> 7

atgatatgaa aaaggaatga acctttctcc cttgcataca aatagggaga aaggttttt 60
tatattaata gattgaggat gagaaatttt ctaaagatgt catattcaaa taggacaacg 120
tcatacacat atagtgtcct gtgttgatt gaaagagctt aataaaattg aaaaggatag 180
gaagtccggg agggatcac ttg gaa aag gtc gat att ttc aag gat atc gct 233
Leu Glu Lys Val Asp Ile Phe Lys Asp Ile Ala
1 5 10
gaa cga aca gga ggc gat ata tac tta gga gtc gta ggt gct gtc cgt 281
Glu Arg Thr Gly Gly Asp Ile Tyr Leu Gly Val Val Gly Ala Val Arg
15 20 25
aca gga aaa tcc acg ttc att aaa aaa ttt atg gag ctt gtg gtg ctc 329
Thr Gly Lys Ser Thr Phe Ile Lys Lys Phe Met Glu Leu Val Val Leu
30 35 40
ccg aat atc agt aac gaa gca gac cgg gcc cga gcg cag gat gaa ctg 377
Pro Asn Ile Ser Asn Glu Ala Asp Arg Ala Arg Ala Gln Asp Glu Leu
45 50 55
ccg cag agc gca gcc ggc aaa acc att atg act aca gag cct aaa ttt 425
Pro Gln Ser Ala Ala Gly Lys Thr Ile Met Thr Thr Glu Pro Lys Phe
60 65 70 75
gtt ccg aat cag gcg atg tct gtt cat gtg tca gac gga ctc gat gtg 473
Val Pro Asn Gln Ala Met Ser Val His Val Ser Asp Gly Leu Asp Val
80 85 90
aat ata aga tta gta gat tgt gta ggt tac aca gtg ccc ggc gct aaa 521
Asn Ile Arg Leu Val Asp Cys Val Gly Tyr Thr Val Pro Gly Ala Lys
95 100 105
gga tat gaa gat gaa aac ggg ccg cgg atg atc aat acg cct tgg tac 569
Gly Tyr Glu Asp Glu Asn Gly Pro Arg Met Ile Asn Thr Pro Trp Tyr
110 115 120
gaa gaa ccg atc cca ttt cat gag gct gct gaa atc ggc aca cga aaa 617
Glu Glu Pro Ile Pro Phe His Glu Ala Ala Glu Ile Gly Thr Arg Lys
125 130 135
gtc att caa gaa cac tcg acc atc gga gtt gtc att acg aca gac ggc 665
Val Ile Gln Glu His Ser Thr Ile Gly Val Val Ile Thr Thr Asp Gly
140 145 150 155
acc att gga gat atc gcc aga agt gac tat ata gag gct gaa gaa aga 713
Thr Ile Gly Asp Ile Ala Arg Ser Asp Tyr Ile Glu Ala Glu Glu Arg
160 165 170
gtc att gaa gag ctg aaa gag gtt ggc aaa cct ttt att atg gtc atc 761
Val Ile Glu Glu Leu Lys Glu Val Gly Lys Pro Phe Ile Met Val Ile
175 180 185
aac tca gtc agg ccg tat cac ccg gaa acg gaa gcc atg cgc cag gat 809
Asn Ser Val Arg Pro Tyr His Pro Glu Thr Glu Ala Met Arg Gln Asp
190 195 200
tta agc gaa aaa tat gat atc ccg gta ttg gca atg agt gta gag agc 857

Leu Ser Glu Lys Tyr Asp Ile Pro Val Leu Ala Met Ser Val Glu Ser			
205	210	215	
atg cgg gaa tca gat gtg ctg agt gtg ctc aga gag gcc ctc tac gag			905
Met Arg Glu Ser Asp Val Leu Ser Val Leu Arg Glu Ala Leu Tyr Glu			
220	225	230	235
ttt ccg gtg cta gaa gtg aat gtc aat ctc cca agc tgg gta atg gtg			953
Phe Pro Val Leu Glu Val Asn Val Asn Leu Pro Ser Trp Val Met Val			
240	245	250	
ctg aaa gaa aac cat tgg ttg cgt gaa agc tat cag gag tcc gtg aag			1001
Leu Lys Glu Asn His Trp Leu Arg Glu Ser Tyr Gln Glu Ser Val Lys			
255	260	265	
gaa acg gtt aag gat att aaa cgg ctc cgg gac gta gac agg gtt gtc			1049
Glu Thr Val Lys Asp Ile Lys Arg Leu Arg Asp Val Asp Arg Val Val			
270	275	280	
ggc caa ttc agc gag ttt gaa ttc att gaa agt gcc gga tta gcc gga			1097
Gly Gln Phe Ser Glu Phe Glu Phe Ile Glu Ser Ala Gly Leu Ala Gly			
285	290	295	
att gag ctg ggc caa ggg gtg gca gaa att gat ttg tac gcg cct gat			1145
Ile Glu Leu Gly Gln Gly Val Ala Glu Ile Asp Leu Tyr Ala Pro Asp			
300	305	310	315
cat cta tat gat caa atc cta aaa gaa gtt gtg ggc gtc gaa atc aga			1193
His Leu Tyr Asp Gln Ile Leu Lys Glu Val Val Gly Val Glu Ile Arg			
320	325	330	
gga aga gac cat ctg ctt gag ctc atg caa gac ttc gcc cat gcg aaa			1241
Gly Arg Asp His Leu Leu Glu Leu Met Gln Asp Phe Ala His Ala Lys			
335	340	345	
aca gaa tat gat caa gtg tct gat gcc tta aaa atg gtc aaa cag acg			1289
Thr Glu Tyr Asp Gln Val Ser Asp Ala Leu Lys Met Val Lys Gln Thr			
350	355	360	
gga tac ggc att gca gcg cct gct tta gct gat atg agt ctc gat gag			1337
Gly Tyr Gly Ile Ala Ala Pro Ala Leu Ala Asp Met Ser Leu Asp Glu			
365	370	375	
ccg gaa att ata agg cag ggc tcg cga ttc ggt gtg agg ctg aaa gct			1385
Pro Glu Ile Ile Arg Gln Gly Ser Arg Phe Gly Val Arg Leu Lys Ala			
380	385	390	395
gtc gct ccg tcg atc cat atg atc aaa gta gat gtc gaa agc gaa ttc			1433
Val Ala Pro Ser Ile His Met Ile Lys Val Asp Val Glu Ser Glu Phe			
400	405	410	
gcc ccg att atc gga acg gaa aaa caa agt gaa gag ctt gta cgc tat			1481
Ala Pro Ile Ile Gly Thr Glu Lys Gln Ser Glu Glu Leu Val Arg Tyr			
415	420	425	
tta atg cag gac ttt gag gat gat ccg ctc tcc atc tgg aat tcc gat			1529
Leu Met Gln Asp Phe Glu Asp Asp Pro Leu Ser Ile Trp Asn Ser Asp			
430	435	440	

H 06291

atc ttc gga agg tcg ctg agc tca att gtg aga gaa ggg att cag gca Ile Phe Gly Arg Ser Leu Ser Ser Ile Val Arg Glu Gly Ile Gln Ala 445 450 455	1577
aag ctg tca ttg atg cct gaa aac gca cgg tat aaa tta aaa gaa aca Lys Leu Ser Leu Met Pro Glu Asn Ala Arg Tyr Lys Leu Lys Glu Thr 460 465 470 475	1625
tta gaa aga atc ata aac gaa ggc tct ggc ggc tta atc gcc atc atc Leu Glu Arg Ile Ile Asn Glu Gly Ser Gly Gly Leu Ile Ala Ile Ile 480 485 490	1673
ctg taa taccggtaga cctctttata gaatgggagg tctttttct ttgctcttaa Leu	1729
taatggaaaa ggatcaagga ataggatgaa aaaaggaaaa aaaggaatat tcgttcggta 1789	
aatcacctta aatccttgac gagcaaggga ttgacgcttt aaaatgcttg atatggcttt 1849	
ttatatgtgt tactctacat acagaaaa	1876

<210> 8
<211> 492
<212> PRT
<213> *Bacillus subtilis*

<400> 8

Leu Glu Lys Val Asp Ile Phe Lys Asp Ile Ala Glu Arg Thr Gly Gly 1 5 10 15	
Asp Ile Tyr Leu Gly Val Val Gly Ala Val Arg Thr Gly Lys Ser Thr 20 25 30	
Phe Ile Lys Phe Met Glu Leu Val Val Leu Pro Asn Ile Ser Asn 35 40 45	
Glu Ala Asp Arg Ala Arg Ala Gln Asp Glu Leu Pro Gln Ser Ala Ala 50 55 60	
Gly Lys Thr Ile Met Thr Thr Glu Pro Lys Phe Val Pro Asn Gln Ala 65 70 75 80	
Met Ser Val His Val Ser Asp Gly Leu Asp Val Asn Ile Arg Leu Val 85 90 95	
Asp Cys Val Gly Tyr Thr Val Pro Gly Ala Lys Gly Tyr Glu Asp Glu 100 105 110	
Asn Gly Pro Arg Met Ile Asn Thr Pro Trp Tyr Glu Glu Pro Ile Pro 115 120 125	
Phe His Glu Ala Ala Glu Ile Gly Thr Arg Lys Val Ile Gln Glu His 130 135 140	
Ser Thr Ile Gly Val Val Ile Thr Thr Asp Gly Thr Ile Gly Asp Ile 145 150 155 160	
Ala Arg Ser Asp Tyr Ile Glu Ala Glu Glu Arg Val Ile Glu Glu Leu 165 170 175	
Lys Glu Val Gly Lys Pro Phe Ile Met Val Ile Asn Ser Val Arg Pro 180 185 190	
Tyr His Pro Glu Thr Glu Ala Met Arg Gln Asp Leu Ser Glu Lys Tyr 195 200 205	
Asp Ile Pro Val Leu Ala Met Ser Val Glu Ser Met Arg Glu Ser Asp 210 215 220	
Val Leu Ser Val Leu Arg Glu Ala Leu Tyr Glu Phe Pro Val Leu Glu 225 230 235 240	
Val Asn Val Asn Leu Pro Ser Trp Val Met Val Leu Lys Glu Asn His	

245	250	255
Trp Leu Arg Glu Ser Tyr Gln Glu Ser Val Lys Glu Thr Val Lys Asp		
260	265	270
Ile Lys Arg Leu Arg Asp Val Asp Arg Val Val Gly Gln Phe Ser Glu		
275	280	285
Phe Glu Phe Ile Glu Ser Ala Gly Leu Ala Gly Ile Glu Leu Gly Gln		
290	295	300
Gly Val Ala Glu Ile Asp Leu Tyr Ala Pro Asp His Leu Tyr Asp Gln		
305	310	315
Ile Leu Lys Glu Val Val Gly Val Glu Ile Arg Gly Arg Asp His Leu		
325	330	335
Leu Glu Leu Met Gln Asp Phe Ala His Ala Lys Thr Glu Tyr Asp Gln		
340	345	350
Val Ser Asp Ala Leu Lys Met Val Lys Gln Thr Gly Tyr Gly Ile Ala		
355	360	365
Ala Pro Ala Leu Ala Asp Met Ser Leu Asp Glu Pro Glu Ile Ile Arg		
370	375	380
Gln Gly Ser Arg Phe Gly Val Arg Leu Lys Ala Val Ala Pro Ser Ile		
385	390	395
His Met Ile Lys Val Asp Val Glu Ser Glu Phe Ala Pro Ile Ile Gly		
405	410	415
Thr Glu Lys Gln Ser Glu Glu Leu Val Arg Tyr Leu Met Gln Asp Phe		
420	425	430
Glu Asp Asp Pro Leu Ser Ile Trp Asn Ser Asp Ile Phe Gly Arg Ser		
435	440	445
Leu Ser Ser Ile Val Arg Glu Gly Ile Gln Ala Lys Leu Ser Leu Met		
450	455	460
Pro Glu Asn Ala Arg Tyr Lys Leu Lys Glu Thr Leu Glu Arg Ile Ile		
465	470	475
Asn Glu Gly Ser Gly Gly Leu Ile Ala Ile Ile Leu		
485	490	

<210> 9
 <211> 1675
 <212> DNA
 <213> Bacillus subtilis

<220>
 <221> CDS
 <222> (201)..(1478)

<220>
 <221> gene
 <222> (1)..(1675)
 <223> spoIVB

<400> 9
 cggatcaagt caaaaacaact gggtaagctg cgcgagaagc gcagcttatt ttttcgtgc 60
 acatccattc gttcatcaagt atatccatcg ttttcttca tatgacagtt ataaataagc 120
 cgtcagaagg caaaaattaaa tgatgttagca gcaagtcata aagaagggtgt gggataggag 180
 cgaggagagt gaagtagtga atg ccc gat aac atc aga aaa gca gta ggt tta 233
 Met Pro Asp Asn Ile Arg Lys Ala Val Gly Leu
 1 5 10

att ctc ctt gtt tcg tta tta agt gta ggt tta tgc aaa ccg cta aaa	281
Ile Leu Leu Val Ser Leu Leu Ser Val Gly Leu Cys Lys Pro Leu Lys	
15 20 25	
gaa tat tta ctg att cca acg caa atg aga gta ttt gaa acc caa aca	329
Glu Tyr Leu Leu Ile Pro Thr Gln Met Arg Val Phe Glu Thr Gln Thr	
30 35 40	
caa gcg att gaa acg agt tta tcg gta aat gct cag aca tca gaa tcc	377
Gln Ala Ile Glu Thr Ser Leu Ser Val Asn Ala Gln Thr Ser Glu Ser	
45 50 55	
tca gaa gcg ttt aca gta aag aaa gat ccg cat gaa atc aag gtg acg	425
Ser Glu Ala Phe Thr Val Lys Lys Asp Pro His Glu Ile Lys Val Thr	
60 65 70 75	
ggc aaa aaa tca ggt gag tca gaa ttg gta tat gat ctt gcc gga ttt	473
Gly Lys Lys Ser Gly Glu Ser Glu Leu Val Tyr Asp Leu Ala Gly Phe	
80 85 90	
cca att aaa aaa aca aaa gtg cat gtt ctt cct gat tta aaa gtt ata	521
Pro Ile Lys Lys Thr Lys Val His Val Leu Pro Asp Leu Lys Val Ile	
95 100 105	
cct ggc gga caa tca atc ggt gta aaa ctt cat tcc gtc ggt gtt ctt	569
Pro Gly Gly Gln Ser Ile Gly Val Lys Leu His Ser Val Gly Val Leu	
110 115 120	
gtc gga ttt cat caa atc aat aca agt gaa ggc aaa aaa tct ccg gga	617
Val Gly Phe His Gln Ile Asn Thr Ser Glu Gly Lys Lys Ser Pro Gly	
125 130 135	
gaa acg gca gga att gaa gcg ggc gac atc att att gag atg aat gga	665
Glu Thr Ala Gly Ile Glu Ala Gly Asp Ile Ile Glu Met Asn Gly	
140 145 150 155	
cag aaa att gaa aaa atg aat gat gta gcc cca ttt att caa aag gct	713
Gln Lys Ile Glu Lys Met Asn Asp Val Ala Pro Phe Ile Gln Lys Ala	
160 165 170	
ggg aaa act ggt gaa tct tta gac tta ctg atc aaa cgt gat aaa cag	761
Gly Lys Thr Gly Glu Ser Leu Asp Leu Leu Ile Lys Arg Asp Lys Gln	
175 180 185	
aaa atc aaa acg aag ctg atc cca gaa aag gat gaa gga gaa ggc aaa	809
Lys Ile Lys Thr Lys Leu Ile Pro Glu Lys Asp Glu Gly Glu Gly Lys	
190 195 200	
tac aga atc ggg tta tat atc aga gat tct gct gct ggc atc ggc act	857
Tyr Arg Ile Gly Leu Tyr Ile Arg Asp Ser Ala Ala Gly Ile Gly Thr	
205 210 215	
atg acc ttt tat gaa ccg aaa aca aaa aaa tac gga gca ctt ggc cac	905
Met Thr Phe Tyr Glu Pro Lys Thr Lys Lys Tyr Gly Ala Leu Gly His	
220 225 230 235	
gtg att tcc gat atg gac aca aag aaa cca att gta gtg gag aat gga	953
Val Ile Ser Asp Met Asp Thr Lys Lys Pro Ile Val Val Glu Asn Gly	
240 245 250	

H 06291

gaa atc gtt aaa tcc act gta aca tca att gaa aaa ggg aca ggc ggt 1001
Glu Ile Val Lys Ser Thr Val Thr Ser Ile Glu Lys Gly Thr Gly Gly
255 260 265

aat ccg gga gaa aaa ctg gcg cga ttt tcc tca gaa cgc aaa acg atc 1049
Asn Pro Gly Glu Lys Leu Ala Arg Phe Ser Ser Glu Arg Lys Thr Ile
270 275 280

ggg gat att aac aga aac agc ccg ttt ggg att ttc ggc aca ctg cat 1097
Gly Asp Ile Asn Arg Asn Ser Pro Phe Gly Ile Phe Gly Thr Leu His
285 290 295

cag ccg att caa aac aac ata tca gat caa gca ttg ccg gtt gcg ttt 1145
Gln Pro Ile Gln Asn Asn Ile Ser Asp Gln Ala Leu Pro Val Ala Phe
300 305 310 315

tct acc gaa gtc aaa aaa ggg ccg gct gaa att tta acg gtt att gat 1193
Ser Thr Glu Val Lys Lys Gly Pro Ala Glu Ile Leu Thr Val Ile Asp
320 325 330

gat gac aaa gta gaa aaa ttc gat att gaa atc gtc agc aca acg ccg 1241
Asp Asp Lys Val Glu Lys Phe Asp Ile Glu Ile Val Ser Thr Thr Pro
335 340 345

caa aaa ttc cct gcg aca aaa gga atg gtg ttg aaa att acc gat cca 1289
Gln Lys Phe Pro Ala Thr Lys Gly Met Val Leu Lys Ile Thr Asp Pro
350 355 360

aga ctg ttg aaa gaa aca gga ggc atc gta cag ggg atg agc gga agc 1337
Arg Leu Leu Lys Glu Thr Gly Ile Val Gln Gly Met Ser Gly Ser
365 370 375

ccg atc att caa aat gga aaa gtg atc ggt gct gtc acc cat gta ttt 1385
Pro Ile Ile Gln Asn Gly Lys Val Ile Gly Ala Val Thr His Val Phe
380 385 390 395

gta aat gac ccg aca agc ggc tac ggt gtt cat att gaa tgg atg ctg 1433
Val Asn Asp Pro Thr Ser Gly Tyr Gly Val His Ile Glu Trp Met Leu
400 405 410

tca gaa gca gga atc gat att tat gga aaa gaa aaa gca agc tga 1478
Ser Glu Ala Gly Ile Asp Ile Tyr Gly Lys Glu Lys Ala Ser
415 420 425

ctgccggagt ttccggcagt ttttttattt tgatccctct tcacttctca gaatacatac 1538

ggtaaaaat acaaaaagaag attttcgac aaattcacgt ttccctgttt gtcaaatttc 1598

attttttagtc gaaaaacaga gaaaaacata gaataacaaa gatatgccac taatattggt 1658

gattatgatt ttttttag 1675

<210> 10
<211> 425
<212> PRT
<213> *Bacillus subtilis*

<400> 10
 Met Pro Asp Asn Ile Arg Lys Ala Val Gly Leu Ile Leu Leu Val Ser
 1 5 10 15
 Leu Leu Ser Val Gly Leu Cys Lys Pro Leu Lys Glu Tyr Leu Leu Ile
 20 25 30
 Pro Thr Gln Met Arg Val Phe Glu Thr Gln Thr Gln Ala Ile Glu Thr
 35 40 45
 Ser Leu Ser Val Asn Ala Gln Thr Ser Glu Ser Ser Glu Ala Phe Thr
 50 55 60
 Val Lys Lys Asp Pro His Glu Ile Lys Val Thr Gly Lys Lys Ser Gly
 65 70 75 80
 Glu Ser Glu Leu Val Tyr Asp Leu Ala Gly Phe Pro Ile Lys Lys Thr
 85 90 95
 Lys Val His Val Leu Pro Asp Leu Lys Val Ile Pro Gly Gly Gln Ser
 100 105 110
 Ile Gly Val Lys Leu His Ser Val Gly Val Leu Val Gly Phe His Gln
 115 120 125
 Ile Asn Thr Ser Glu Gly Lys Lys Ser Pro Gly Glu Thr Ala Gly Ile
 130 135 140
 Glu Ala Gly Asp Ile Ile Glu Met Asn Gly Gln Lys Ile Glu Lys
 145 150 155 160
 Met Asn Asp Val Ala Pro Phe Ile Gln Lys Ala Gly Lys Thr Gly Glu
 165 170 175
 Ser Leu Asp Leu Leu Ile Lys Arg Asp Lys Gln Lys Ile Lys Thr Lys
 180 185 190
 Leu Ile Pro Glu Lys Asp Glu Gly Glu Gly Lys Tyr Arg Ile Gly Leu
 195 200 205
 Tyr Ile Arg Asp Ser Ala Ala Gly Ile Gly Thr Met Thr Phe Tyr Glu
 210 215 220
 Pro Lys Thr Lys Tyr Gly Ala Leu Gly His Val Ile Ser Asp Met
 225 230 235 240
 Asp Thr Lys Pro Ile Val Val Glu Asn Gly Glu Ile Val Lys Ser
 245 250 255
 Thr Val Thr Ser Ile Glu Lys Gly Thr Gly Gly Asn Pro Gly Glu Lys
 260 265 270
 Leu Ala Arg Phe Ser Ser Glu Arg Lys Thr Ile Gly Asp Ile Asn Arg
 275 280 285
 Asn Ser Pro Phe Gly Ile Phe Gly Thr Leu His Gln Pro Ile Gln Asn
 290 295 300
 Asn Ile Ser Asp Gln Ala Leu Pro Val Ala Phe Ser Thr Glu Val Lys
 305 310 315 320
 Lys Gly Pro Ala Glu Ile Leu Thr Val Ile Asp Asp Asp Lys Val Glu
 325 330 335
 Lys Phe Asp Ile Glu Ile Val Ser Thr Thr Pro Gln Lys Phe Pro Ala
 340 345 350
 Thr Lys Gly Met Val Leu Lys Ile Thr Asp Pro Arg Leu Leu Lys Glu
 355 360 365
 Thr Gly Gly Ile Val Gln Gly Met Ser Gly Ser Pro Ile Ile Gln Asn
 370 375 380
 Gly Lys Val Ile Gly Ala Val Thr His Val Phe Val Asn Asp Pro Thr
 385 390 395 400
 Ser Gly Tyr Gly Val His Ile Glu Trp Met Leu Ser Glu Ala Gly Ile
 405 410 415
 Asp Ile Tyr Gly Lys Glu Lys Ala Ser
 420 425

H 06291

<211> 1900
<212> DNA
<213> *Bacillus subtilis*

<220>
<221> CDS
<222> (201)..(1703)

<220>
<221> gene
<222> (1)..(1900)
<223> spoIVCA

<220>
<221> misc_feature
<222> (201)..(203)
<223> First codon translated as Met.

<400> 11
ttttgcata tcattgaaac gtttaataac actatagttt aatttaaaat ctccctcattt 60
ggacaaacag ctgttacata gcattaccca aggggtgatg cattttatga aagtgataat 120
catcgagggc cccgaagctg acaaatgcat taacgattgc tatcattatt taataaaaact 180
ttataggaag gagattcagg gtg ata gca ata tat gta agg gta tcg acc gag 233
Val Ile Ala Ile Tyr Val Arg Val Ser Thr Glu
1 5 10
gaa caa gcg atc aag gga tcg agc atc gac agc caa atc gag gcc tgt 281
Glu Gln Ala Ile Lys Gly Ser Ser Ile Asp Ser Gln Ile Glu Ala Cys
15 20 25
ata aag aaa gca ggg act aaa gat gtg ctg aag tat gca gat gaa gga 329
Ile Lys Lys Ala Gly Thr Lys Asp Val Leu Lys Tyr Ala Asp Glu Gly
30 35 40
ttt tca gga gag ctt tta gaa cgt ccg gct ttg aat cgc ttg agg gag 377
Phe Ser Gly Glu Leu Leu Glu Arg Pro Ala Leu Asn Arg Leu Arg Glu
45 50 55
gat gca agc aag gga ctt ata agt caa gtc att tgt tac gat cct gac 425
Asp Ala Ser Lys Gly Leu Ile Ser Gln Val Ile Cys Tyr Asp Pro Asp
60 65 70 75
cgt ctt tct cgg aaa tta atg aat cag cta atc att gat gac gaa ttg 473
Arg Leu Ser Arg Lys Leu Met Asn Gln Leu Ile Ile Asp Asp Glu Leu
80 85 90
cga aag cga aac ata cct ttg att ttt gta aat ggt gaa tac gcc aat 521
Arg Lys Arg Asn Ile Pro Leu Ile Phe Val Asn Gly Glu Tyr Ala Asn
95 100 105
tct cca gaa ggt caa ttg ttt ttc gca atg cgc ggg gca atc tca gaa 569
Ser Pro Glu Gly Gln Leu Phe Phe Ala Met Arg Gly Ala Ile Ser Glu
110 115 120
ttt gaa aaa gcc aaa atc aaa gaa cgg aca tca agc ggc cga ctt caa 617
Phe Glu Lys Ala Lys Glu Arg Thr Ser Ser Gly Arg Leu Gln

	125	130	135	
aaa atg aaa aaa ggc atg atc att aaa gat tct aaa cta tat ggc tat				665
Lys Met Lys Lys Gly Met Ile Ile Lys Asp Ser Lys Leu Tyr Gly Tyr				
140	145	150	155	
aaa ttt gtt aaa gag aaa aga act ctt gag ata tta gaa gag gaa gca				713
Lys Phe Val Lys Glu Lys Arg Thr Leu Glu Ile Leu Glu Glu Ala				
160	165	170		
aaa atc att cgg atg att ttt aac tat ttc acc gat cat aaa agc cct				761
Lys Ile Ile Arg Met Ile Phe Asn Tyr Phe Thr Asp His Lys Ser Pro				
175	180	185		
ttt ttc ggc aga gta aat ggt att gct cta cat tta act cag atg ggg				809
Phe Phe Gly Arg Val Asn Gly Ile Ala Leu His Leu Thr Gln Met Gly				
190	195	200		
gtt aaa aca aaa aaa ggc gcc aaa gta tgg cac agg cag gtt gtt cgg				857
Val Lys Thr Lys Lys Gly Ala Lys Val Trp His Arg Gln Val Val Arg				
205	210	215		
caa ata tta atg aac tct tcc tat aag ggt gaa cat aga cag tat aaa				905
Gln Ile Leu Met Asn Ser Ser Tyr Lys Gly Glu His Arg Gln Tyr Lys				
220	225	230	235	
tat gat aca gag ggt tcc tat gtt tca aag cag gca ggg aac aaa tct				953
Tyr Asp Thr Glu Gly Ser Tyr Val Ser Lys Gln Ala Gly Asn Lys Ser				
240	245	250		
ata att aaa ata agg cct gaa gaa gaa caa atc act gtg aca att cca				1001
Ile Ile Lys Ile Arg Pro Glu Glu Gln Ile Thr Val Thr Ile Pro				
255	260	265		
gca att gtt cca gct gaa caa tgg gat tat gct caa gaa ctc tta ggt				1049
Ala Ile Val Pro Ala Glu Gln Trp Asp Tyr Ala Gln Glu Leu Leu Gly				
270	275	280		
caa agt aaa aga aaa cac ttg agt atc agc cct cac aat tac ttg tta				1097
Gln Ser Lys Arg Lys His Leu Ser Ile Ser Pro His Asn Tyr Leu Leu				
285	290	295		
tcg ggt ttg gtt aga tgc gga aaa tgc gga aat acc atg aca ggg aag				1145
Ser Gly Leu Val Arg Cys Gly Lys Cys Gly Asn Thr Met Thr Gly Lys				
300	305	310	315	
aaa aga aaa tca cat ggt aaa gac tac tat gta tat act tgc cgg aaa				1193
Lys Arg Lys Ser His Gly Lys Asp Tyr Tyr Val Tyr Thr Cys Arg Lys				
320	325	330		
aat tat tct ggc gca aag gac cgc ggc tgc gga aaa gaa atg tct gag				1241
Asn Tyr Ser Gly Ala Lys Asp Arg Gly Cys Gly Lys Glu Met Ser Glu				
335	340	345		
aat aaa ttg aac cgg cat gta tgg ggt gaa att ttt aaa ttc atc aca				1289
Asn Lys Leu Asn Arg His Val Trp Gly Glu Ile Phe Lys Phe Ile Thr				
350	355	360		
aat cct caa aag tat gtt tct ttt aaa gag gct gaa caa tca aat cac				1337

Asn Pro Gln Lys Tyr Val Ser Phe Lys Glu Ala Glu Gln Ser Asn His		
365	370	375
ctg tct gat gaa tta gaa ctt att gaa aaa gag ata gag aaa aca aaa		1385
Leu Ser Asp Glu Leu Glu Leu Ile Glu Lys Glu Ile Glu Lys Thr Lys		
380	385	390
aaa ggc cgc aag cgt ctt tta acg cta atc agc cta agc gat gac gat		1433
Lys Gly Arg Lys Arg Leu Leu Thr Leu Ile Ser Leu Ser Asp Asp Asp		
400	405	410
gat tta gac ata gat gaa atc aaa gca caa att att gaa ctg caa aaa		1481
Asp Leu Asp Ile Asp Glu Ile Lys Ala Gln Ile Ile Glu Leu Gln Lys		
415	420	425
aag caa aat cag ctt act gaa aag tgt aac aga atc cag tca aaa atg		1529
Lys Gln Asn Gln Leu Thr Glu Lys Cys Asn Arg Ile Gln Ser Lys Met		
430	435	440
aaa gtc cta gat gat acg agc tca agt gaa aat gct cta aaa aga gcc		1577
Lys Val Leu Asp Asp Thr Ser Ser Glu Asn Ala Leu Lys Arg Ala		
445	450	455
atc gac tat ttt caa tca atc ggt gca gat aac tta act ctt gaa gat		1625
Ile Asp Tyr Phe Gln Ser Ile Gly Ala Asp Asn Leu Thr Leu Glu Asp		
460	465	470
aaa aaa aca att gtt aac ttt atc gtg aaa gaa gtt acc att gtg gat		1673
Lys Lys Thr Ile Val Asn Phe Ile Val Lys Glu Val Thr Ile Val Asp		
480	485	490
tct gac acc ata tat att gaa acg tat taa agaggggtgt atgcacccccc		1723
Ser Asp Thr Ile Tyr Ile Glu Thr Tyr		
495	500	
cttttgtaat tacaatctca ttttcaatac acctcgctgc atacgtcgcc acctttgtcc		1783
cttttccagc ggaatagctt tcaattcctt taataagccc gatcggtccg atggagatta		1843
agtccctctgc atcctcacct gtatttcga acttttcac aatatggcg accaagc		1900
<210> 12		
<211> 500		
<212> PRT		
<213> Bacillus subtilis		
<400> 12		
Val Ile Ala Ile Tyr Val Arg Val Ser Thr Glu Glu Gln Ala Ile Lys		
1	5	10
Gly Ser Ser Ile Asp Ser Gln Ile Glu Ala Cys Ile Lys Lys Ala Gly		
20	25	30
Thr Lys Asp Val Leu Lys Tyr Ala Asp Glu Gly Phe Ser Gly Glu Leu		
35	40	45
Leu Glu Arg Pro Ala Leu Asn Arg Leu Arg Glu Asp Ala Ser Lys Gly		
50	55	60
Leu Ile Ser Gln Val Ile Cys Tyr Asp Pro Asp Arg Leu Ser Arg Lys		
65	70	75
Leu Met Asn Gln Leu Ile Asp Asp Glu Leu Arg Lys Arg Asn Ile		80

85	90	95
Pro Leu Ile Phe Val Asn Gly Glu	Tyr Ala Asn Ser Pro	Glu Gly Gln
100	105	110
Leu Phe Phe Ala Met Arg Gly Ala	Ile Ser Glu Phe Glu	Lys Ala Lys
115	120	125
Ile Lys Glu Arg Thr Ser Ser Gly Arg	Leu Gln Lys Met Lys Lys Gly	Gly
130	135	140
Met Ile Ile Lys Asp Ser Lys Leu Tyr Gly	Tyr Lys Phe Val Lys Glu	160
145	150	155
Lys Arg Thr Leu Glu Ile Leu Glu Glu	Ala Lys Ile Ile Arg Met	
165	170	175
Ile Phe Asn Tyr Phe Thr Asp His Lys	Ser Pro Phe Phe Gly Arg Val	
180	185	190
Asn Gly Ile Ala Leu His Leu Thr Gln	Met Gly Val Lys Thr Lys Lys	
195	200	205
Gly Ala Lys Val Trp His Arg Gln Val	Val Arg Gln Ile Leu Met Asn	
210	215	220
Ser Ser Tyr Lys Gly Glu His Arg Gln	Tyr Lys Tyr Asp Thr Glu Gly	240
225	230	235
Ser Tyr Val Ser Lys Gln Ala Gly Asn	Lys Ser Ile Ile Lys Ile Arg	
245	250	255
Pro Glu Glu Glu Gln Ile Thr Val Thr	Ile Pro Ala Ile Val Pro Ala	
260	265	270
Glu Gln Trp Asp Tyr Ala Gln Glu	Leu Leu Gly Gln Ser Lys Arg Lys	
275	280	285
His Leu Ser Ile Ser Pro His Asn Tyr	Leu Ser Gly Leu Val Arg	
290	295	300
Cys Gly Lys Cys Gly Asn Thr Met Thr	Gly Lys Arg Lys Ser His	
305	310	315
Gly Lys Asp Tyr Tyr Val Tyr Thr Cys	Arg Lys Asn Tyr Ser Gly Ala	
325	330	335
Lys Asp Arg Gly Cys Gly Lys Glu	Met Ser Glu Asn Lys Leu Asn Arg	
340	345	350
His Val Trp Gly Glu Ile Phe Lys	Phe Ile Thr Asn Pro Gln Lys Tyr	
355	360	365
Val Ser Phe Lys Glu Ala Glu Gln	Ser Asn His Leu Ser Asp Glu Leu	
370	375	380
Glu Leu Ile Glu Lys Glu Ile Glu	Lys Thr Lys Gly Arg Lys Arg	
385	390	395
Leu Leu Thr Leu Ile Ser Leu Ser Asp	Asp Asp Asp Leu Asp Ile Asp	
405	410	415
Glu Ile Lys Ala Gln Ile Ile Glu	Leu Gln Lys Lys Gln Asn Gln Leu	
420	425	430
Thr Glu Lys Cys Asn Arg Ile Gln	Ser Lys Met Lys Val Leu Asp Asp	
435	440	445
Thr Ser Ser Glu Asn Ala Leu Lys	Arg Ala Ile Asp Tyr Phe Gln	
450	455	460
Ser Ile Gly Ala Asp Asn Leu Thr	Leu Glu Asp Lys Lys Thr Ile Val	
465	470	475
Asn Phe Ile Val Lys Glu Val Thr Ile	Val Asp Ser Asp Thr Ile Tyr	
485	490	495
Ile Glu Thr Tyr		
500		

<210> 13

<211> 868

<212> DNA

H 06291

<213> Bacillus subtilis

<220>

<221> CDS

<222> (201)..(671)

<220>

<221> gene

<222> (1)..(868)

<223> spoIVCB

<400> 13

ttcatccccca tccccccata cctttgttca tttcaatgta tggcgcttg atgaagaata 60

tttttaacat ttgaagtttag tatgtgtcacc accaaagccg gactcccccg cgagaaaattt 120

cccggtacag acacagacag cctcccggtc acatacattt acatataggc tttgcctac 180

atacttttgt ggaggtgacg atg gtg aca ggt gtt ttc gca gcg ctc ggc ttt 233
Met Val Thr Gly Val Phe Ala Ala Leu Gly Phe
1 5 10

gtt gtt aaa gag ctt gtc ttt tta gta tct tac gtg aaa aac aat gcc 281
Val Val Lys Glu Leu Val Phe Leu Val Ser Tyr Val Lys Asn Asn Ala
15 20 25

ttt cca caa ccg ctc tca agc agc gaa gaa aaa aaa tac tta gag ctc 329
Phe Pro Gln Pro Leu Ser Ser Glu Glu Lys Lys Tyr Leu Glu Leu
30 35 40

atg gct aaa ggg gat gaa cat gcc aga aac atg ctg att gag cat aat 377
Met Ala Lys Gly Asp Glu His Ala Arg Asn Met Leu Ile Glu His Asn
45 50 55

ctt cgc ttg gtc gcc cat att gtg aaa aag ttc gaa aat aca ggt gag 425
Leu Arg Leu Val Ala His Ile Val Lys Lys Phe Glu Asn Thr Gly Glu
60 65 70 75

gat gca gag gac tta atc tcc atc gga acg atc ggg ctt att aaa gga 473
Asp Ala Glu Asp Leu Ile Ser Ile Gly Thr Ile Gly Leu Ile Lys Gly
80 85 90

att gaa agc tat tcc gct gga aaa ggg aca aag gtg gcg acg tat gca 521
Ile Glu Ser Tyr Ser Ala Gly Lys Gly Thr Lys Val Ala Thr Tyr Ala
95 100 105

gcg agg tgt att gaa aat gag att gta att aca aaa ggg ggg tgc ata 569
Ala Arg Cys Ile Glu Asn Glu Ile Val Ile Thr Lys Gly Gly Cys Ile
110 115 120

cac ccc tct tta ata cgt ttc aat ata tat ggt gtc aga atc cac aat 617
His Pro Ser Leu Ile Arg Phe Asn Ile Tyr Gly Val Arg Ile His Asn
125 130 135

ggg aac ttc ttt cac gat aaa gtt aac aat tgt ttt ttt atc ttc aag 665
Gly Asn Phe Phe His Asp Lys Val Asn Asn Cys Phe Phe Ile Phe Lys
140 145 150 155

agt taa gttatctgca ccgattgatt gaaaaatagtc gatggctt ttttagagcat 721

H 06291

Ser

tttcacttga gctcgatca tctaggactt tcattttga ctggattctg ttacactttt 781
cagtaagctg attttgctt ttttgcagtt caataatgg tgctttgatt tcatctatgt 841
ctaaatcatc gtcatcgctt aggctga 868

<210> 14
<211> 156
<212> PRT
<213> Bacillus subtilis

<400> 14
Met Val Thr Gly Val Phe Ala Ala Leu Gly Phe Val Val Lys Glu Leu
1 5 10 15
Val Phe Leu Val Ser Tyr Val Lys Asn Asn Ala Phe Pro Gln Pro Leu
20 25 30
Ser Ser Ser Glu Glu Lys Lys Tyr Leu Glu Leu Met Ala Lys Gly Asp
35 40 45
Glu His Ala Arg Asn Met Leu Ile Glu His Asn Leu Arg Leu Val Ala
50 55 60
His Ile Val Lys Lys Phe Glu Asn Thr Gly Glu Asp Ala Glu Asp Leu
65 70 75 80
Ile Ser Ile Gly Thr Ile Gly Leu Ile Lys Gly Ile Glu Ser Tyr Ser
85 90 95
Ala Gly Lys Gly Thr Lys Val Ala Thr Tyr Ala Ala Arg Cys Ile Glu
100 105 110
Asn Glu Ile Val Ile Thr Lys Gly Gly Cys Ile His Pro Ser Leu Ile
115 120 125
Arg Phe Asn Ile Tyr Gly Val Arg Ile His Asn Gly Asn Phe Phe His
130 135 140
Asp Lys Val Asn Asn Cys Phe Phe Ile Phe Lys Ser
145 150 155

<210> 15
<211> 1192
<212> DNA
<213> Bacillus subtilis

<220>
<221> CDS
<222> (201)..(995)

<220>
<221> gene
<222> (1)..(1192)
<223> spoIVFA

<400> 15
acaaaggaat gatggctaaat ttttcggagt aagatcttaa tgtgatagaa 60
tcaaagagaa gaatctgaca aagcatatgc tgtgtcaggt tttttttgt ttttgcctgc 120
tttgttcttg actaaaccga atatttgcca tggacaagac atatgatgta caaacccaac 180

H 06291

gaatgcaaag	gatgatggca	atg	agt	cac	aga	gca	gat	gaa	atc	aga	aaa	cga	233			
		Met	Ser	His	Arg	Ala	Asp	Glu	Ile	Arg	Lys	Arg				
		1							5				10			
tta	gag	aaa	aga	aga	aag	cag	ctt	tcc	ggc	tca	aaa	cgt	ttc	tct	act	281
Leu	Glu	Lys	Arg	Arg	Lys	Gln	Leu	Ser	Gly	Ser	Lys	Arg	Phe	Ser	Thr	
		15						20					25			
cag	aca	gtt	tct	gaa	aag	cag	aaa	ccc	ccg	tcc	tgg	gtg	atg	gta	acg	329
Gln	Thr	Val	Ser	Glu	Lys	Gln	Lys	Pro	Pro	Ser	Trp	Val	Met	Val	Thr	
		30					35					40				
gat	cag	gaa	aag	cat	gga	aca	ctt	ccg	gtc	tac	gaa	gat	aac	atg	cca	377
Asp	Gln	Glu	Lys	His	Gly	Thr	Leu	Pro	Val	Tyr	Glu	Asp	Asn	Met	Pro	
		45					50				55					
aca	ttc	aac	gga	aaa	cac	cca	ttg	gtg	aaa	aca	gat	tca	att	atc	ctg	425
Thr	Phe	Asn	Gly	Lys	His	Pro	Leu	Val	Lys	Thr	Asp	Ser	Ile	Ile	Leu	
		60					65				70			75		
aaa	tgt	ctt	ctg	tcg	gcc	tgc	ctt	gtt	ctc	gtt	tca	gct	ata	gcc	tat	473
Lys	Cys	Leu	Leu	Ser	Ala	Cys	Leu	Val	Leu	Val	Ser	Ala	Ile	Ala	Tyr	
			80					85					90			
aaa	aca	aac	att	gga	ccc	gtc	agt	cag	att	aaa	ccc	gcc	gta	gcc	aaa	521
Lys	Thr	Asn	Ile	Gly	Pro	Val	Ser	Gln	Ile	Lys	Pro	Ala	Val	Ala	Lys	
		95						100					105			
acc	ttt	gaa	act	gaa	ttt	caa	ttt	gct	tca	gca	agc	cat	tgg	ttc	gaa	569
Thr	Phe	Glu	Thr	Glu	Phe	Gln	Phe	Ala	Ser	Ala	Ser	His	Trp	Phe	Glu	
		110					115					120				
acc	aaa	ttc	gga	aat	ccg	ctt	gct	ttc	ctg	gct	cct	gaa	cac	aaa	aat	617
Thr	Lys	Phe	Gly	Asn	Pro	Leu	Ala	Phe	Leu	Ala	Pro	Glu	His	Lys	Asn	
		125					130				135					
aag	gaa	cag	cag	att	gaa	gta	ggc	aaa	gat	ctg	atc	gcg	cct	gca	tcc	665
Lys	Glu	Gln	Gln	Ile	Glu	Val	Gly	Lys	Asp	Leu	Ile	Ala	Pro	Ala	Ser	
		140				145				150			155			
ggg	aaa	gta	cag	cag	gat	ttt	cag	gac	aat	ggg	gaa	gga	att	aaa	gtc	713
Gly	Lys	Val	Gln	Gln	Asp	Phe	Gln	Asp	Asn	Gly	Glu	Gly	Ile	Lys	Val	
		160					165					170				
gaa	aca	agc	agt	gat	aag	att	gat	agc	gta	aaa	gaa	ggc	tat	gtg	gtt	761
Glu	Thr	Ser	Ser	Asp	Lys	Ile	Asp	Ser	Val	Lys	Glu	Gly	Tyr	Val	Val	
		175					180				185					
gaa	gtc	agc	aaa	gac	agc	caa	acg	gga	ctg	acg	gtt	aag	gtg	cag	cat	809
Glu	Val	Ser	Lys	Asp	Ser	Gln	Thr	Gly	Leu	Thr	Val	Lys	Val	Gln	His	
		190					195				200					
gct	gac	acc	acc	tat	agt	atc	tat	ggc	gag	ctc	aaa	gat	gtg	gat	gtt	857
Ala	Asp	Asn	Thr	Tyr	Ser	Ile	Tyr	Gly	Glu	Leu	Lys	Asp	Val	Asp	Val	
		205					210				215					
gct	tta	tat	gat	ttt	gtg	gat	aaa	ggc	aaa	aag	ctc	ggt	tcg	att	aag	905
Ala	Leu	Tyr	Asp	Phe	Val	Asp	Lys	Gly	Lys	Lys	Leu	Gly	Ser	Ile	Lys	
		220				225				230			235			

ctt gat gat cat aat aaa ggg gtc tat tat ttt gcc atg aaa gac ggc 953
Leu Asp Asp His Asn Lys Gly Val Tyr Tyr Phe Ala Met Lys Asp Gly
240 245 250

gat aaa ttt att gat ccg att cag gtg att tca ttt gaa taa 995
Asp Lys Phe Ile Asp Pro Ile Gln Val Ile Ser Phe Glu
255 260 265

atggctcgac cttatcttaa agatccatgt gcattttttt ctttggatta ttgcggcgct 1055
gggcttgctc acaggccata tgaaagcatt attatgtctg ctcctgattt tattgattca 1115
tgagctgggg catgctgctc tggctgtgtt ttttcttgg agaatcaagc gtgtttttt 1175
gctgccgtt ggcggaa 1192

<210> 16
<211> 264
<212> PRT
<213> *Bacillus subtilis*

<400> 16

Met Ser His Arg Ala Asp Glu Ile Arg Lys Arg Leu Glu Lys Arg Arg
1 5 10 15
Lys Gln Leu Ser Gly Ser Lys Arg Phe Ser Thr Gln Thr Val Ser Glu
20 25 30
Lys Gln Lys Pro Pro Ser Trp Val Met Val Thr Asp Gln Glu Lys His
35 40 45
Gly Thr Leu Pro Val Tyr Glu Asp Asn Met Pro Thr Phe Asn Gly Lys
50 55 60
His Pro Leu Val Lys Thr Asp Ser Ile Ile Leu Lys Cys Leu Leu Ser
65 70 75 80
Ala Cys Leu Val Leu Val Ser Ala Ile Ala Tyr Lys Thr Asn Ile Gly
85 90 95
Pro Val Ser Gln Ile Lys Pro Ala Val Ala Lys Thr Phe Glu Thr Glu
100 105 110
Phe Gln Phe Ala Ser Ala Ser His Trp Phe Glu Thr Lys Phe Gly Asn
115 120 125
Pro Leu Ala Phe Leu Ala Pro Glu His Lys Asn Lys Glu Gln Gln Ile
130 135 140
Glu Val Gly Lys Asp Leu Ile Ala Pro Ala Ser Gly Lys Val Gln Gln
145 150 155 160
Asp Phe Gln Asp Asn Gly Glu Gly Ile Lys Val Glu Thr Ser Ser Asp
165 170 175
Lys Ile Asp Ser Val Lys Glu Gly Tyr Val Val Glu Val Ser Lys Asp
180 185 190
Ser Gln Thr Gly Leu Thr Val Lys Val Gln His Ala Asp Asn Thr Tyr
195 200 205
Ser Ile Tyr Gly Glu Leu Lys Asp Val Asp Val Ala Leu Tyr Asp Phe
210 215 220
Val Asp Lys Gly Lys Lys Leu Gly Ser Ile Lys Leu Asp Asp His Asn
225 230 235 240
Lys Gly Val Tyr Tyr Phe Ala Met Lys Asp Gly Asp Lys Phe Ile Asp
245 250 255
Pro Ile Gln Val Ile Ser Phe Glu
260

<210> 17
<211> 1264
<212> DNA
<213> *Bacillus subtilis*

<220>
<221> CDS
<222> (201)..(1067)

<220>
<221> gene
<222> (1)..(1264)
<223> *spoIVFB*

<220>
<221> misc_feature
<222> (201)..(203)
<223> First codon translated as Met.

<400> 17
actgacgggtt aaggtgcagc atgctgacaa cacctatagt atctatggcg agctcaaaga 60
tgtggatgtt gctttatataatg attttgtgga taaaggcaaa aagctcggtt cgattaagct 120
tcatgtatcat aataaaagggg tctattattt tgccatgaaa gacggcgata aatttattga 180
tccgatttcag gtgatttcat ttg aat aaa tgg ctc gac ctt atc tta aag atc 233
Leu Asn Lys Trp Leu Asp Leu Ile Leu Lys Ile
1 5 10

cat gtg cat cct ttt ctt tgg att att gcg gcg ctg ggc ttg ctc aca 281
His Val His Pro Phe Leu Trp Ile Ile Ala Ala Leu Gly Leu Leu Thr
15 20 25

ggc cat atg aaa gca tta tta tgt ctg ctc ctg att gta ttg att cat 329
Gly His Met Lys Ala Leu Leu Cys Leu Leu Ile Val Leu Ile His
30 35 40

gag ctg ggg cat gct gct ctg gct gtg ttt ttt tct tgg aga atc aag 377
Glu Leu Gly His Ala Ala Leu Ala Val Phe Phe Ser Trp Arg Ile Lys
45 50 55

cgt gtt ttt ttg ctg ccg ttt ggc gga acg gtc gaa gtg gaa gag cac 425
Arg Val Phe Leu Leu Pro Phe Gly Gly Thr Val Glu Val Glu Glu His
60 65 70 75

ggg aat cgg ccg tta aag gaa gag ttt gcg gtc att att gcc gga cct 473
Gly Asn Arg Pro Leu Lys Glu Glu Phe Ala Val Ile Ile Ala Gly Pro
80 85 90

ctt cag cac atc tgg ctt cag ttt gcc gcc tgg atg ctt gca gaa gtc 521
Leu Gln His Ile Trp Leu Gln Phe Ala Ala Trp Met Leu Ala Glu Val
95 100 105

tca gtg att cat cag cat acc ttt gaa ctc ttc acc ttt tat aat ctt 569
Ser Val Ile His Gln His Thr Phe Glu Leu Phe Thr Phe Tyr Asn Leu
110 115 120

tct atc tta ttt gtc aat tta ctg ccg atc tgg ccg ctg gat gga gga	617
Ser Ile Leu Phe Val Asn Leu Leu Pro Ile Trp Pro Leu Asp Gly Gly	
125 130 135	
aaa ctg tta ttt ttg ttg ttt tcc aaa cag ctg cct ttt caa aag gct	665
Lys Leu Leu Phe Leu Leu Phe Ser Lys Gln Leu Pro Phe Gln Lys Ala	
140 145 150 155	
cac cgg ctt aat cta aaa acg tcg ctc tgc ttc tgc ctg ctg ctc ggg	713
His Arg Leu Asn Leu Lys Thr Ser Leu Cys Phe Cys Leu Leu Gly	
160 165 170	
tgc tgg gtt tta ttc gtg att cct ctg caa atc agc gca tgg gtt ttg	761
Cys Trp Val Leu Phe Val Ile Pro Leu Gln Ile Ser Ala Trp Val Leu	
175 180 185	
ttt gtc ttt ctg gct gtt tcc ttg ttt gag gaa tat cgg caa agg cac	809
Phe Val Phe Leu Ala Val Ser Leu Phe Glu Glu Tyr Arg Gln Arg His	
190 195 200	
tat atc cat gtg aga ttt ctc ctc gaa agg tat tac gga aaa aac agg	857
Tyr Ile His Val Arg Phe Leu Leu Glu Arg Tyr Tyr Gly Lys Asn Arg	
205 210 215	
gag ctt gag aaa ctt ctg ccg ctg aca gta aag gcg gag gat aaa gtc	905
Glu Leu Glu Lys Leu Leu Pro Leu Thr Val Lys Ala Glu Asp Lys Val	
220 225 230 235	
tat cat gtg atg gcc gag ttc aaa cgt ggc tgt aag cat ccg att att	953
Tyr His Val Met Ala Glu Phe Lys Arg Gly Cys Lys His Pro Ile Ile	
240 245 250	
ata gaa aaa tca ggc caa aag ctc agc cag ctt gac gag aat gaa gtg	1001
Ile Glu Lys Ser Gly Gln Lys Leu Ser Gln Leu Asp Glu Asn Glu Val	
255 260 265	
ctg cac gct tac ttt gcc gat aag cgg acg aat tct tcc atg gag gaa	1049
Leu His Ala Tyr Phe Ala Asp Lys Arg Thr Asn Ser Ser Met Glu Glu	
270 275 280	
ctg ctt ctg ccc tac taa aactgattga caaacgcctt gtattttgt	1097
Leu Leu Leu Pro Tyr	
285	
atattttta atgttatgga tgttagcacca ttgctacaac cgctcagttac aggtgttaag	1157
agcttttaca gccccctggatctggcag tcttagtcta ataggaggtg cagagaatgt	1217
acgcaatcat taaaacaggc ggtaaaacaaa tcaaaggta agaaggc	1264
<210> 18	
<211> 288	
<212> PRT	
<213> <i>Bacillus subtilis</i>	
<400> 18	
Leu Asn Lys Trp Leu Asp Leu Ile Leu Lys Ile His Val His Pro Phe	

1	5	10	15
Leu Trp Ile Ile Ala Ala Leu Gly Leu Leu Thr Gly His Met Lys Ala			
20	25	30	
Leu Leu Cys Leu Leu Leu Ile Val Leu Ile His Glu Leu Gly His Ala			
35	40	45	
Ala Leu Ala Val Phe Phe Ser Trp Arg Ile Lys Arg Val Phe Leu Leu			
50	55	60	
Pro Phe Gly Gly Thr Val Glu Val Glu His Gly Asn Arg Pro Leu			
65	70	75	80
Lys Glu Glu Phe Ala Val Ile Ile Ala Gly Pro Leu Gln His Ile Trp			
85	90	95	
Leu Gln Phe Ala Ala Trp Met Leu Ala Glu Val Ser Val Ile His Gln			
100	105	110	
His Thr Phe Glu Leu Phe Thr Phe Tyr Asn Leu Ser Ile Leu Phe Val			
115	120	125	
Asn Leu Leu Pro Ile Trp Pro Leu Asp Gly Gly Lys Leu Leu Phe Leu			
130	135	140	
Leu Phe Ser Lys Gln Leu Pro Phe Gln Lys Ala His Arg Leu Asn Leu			
145	150	155	160
Lys Thr Ser Leu Cys Phe Cys Leu Leu Leu Gly Cys Trp Val Leu Phe			
165	170	175	
Val Ile Pro Leu Gln Ile Ser Ala Trp Val Leu Phe Val Phe Leu Ala			
180	185	190	
Val Ser Leu Phe Glu Glu Tyr Arg Gln Arg His Tyr Ile His Val Arg			
195	200	205	
Phe Leu Leu Glu Arg Tyr Tyr Gly Lys Asn Arg Glu Leu Glu Lys Leu			
210	215	220	
Leu Pro Leu Thr Val Lys Ala Glu Asp Lys Val Tyr His Val Met Ala			
225	230	235	240
Glu Phe Lys Arg Gly Cys Lys His Pro Ile Ile Ile Glu Lys Ser Gly			
245	250	255	
Gln Lys Leu Ser Gln Leu Asp Glu Asn Glu Val Leu His Ala Tyr Phe			
260	265	270	
Ala Asp Lys Arg Thr Asn Ser Ser Met Glu Glu Leu Leu Leu Pro Tyr			
275	280	285	